

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2003, 23:43:34 ; Search time 36 Seconds

(without alignments)
958.564 Million cell updates/sec

Title: US-10-025-380-1081

Perfect score: 4321
Sequence: 1 MIIQAHLSLCLMLYLATG.....DKGDNVESQAASVVKPLRS 832

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4321	100.0	832	1	CADH_HUMAN
2	3450	79.8	827	1	CADH_MOUSE
3	3414	79.0	827	1	CADH_RAT
4	954	22.1	829	1	CADG_RABIT
5	922.5	21.3	829	1	CADG_HUMAN
6	877.5	20.3	830	1	CADG_MOUSE
7	715	16.5	913	1	CAD4_CHICK
8	705	16.3	913	1	CAD4_MOUSE
9	699.5	16.2	906	1	CAD2_HUMAN
10	694.5	16.1	912	1	CAD2_CHICK
11	688.5	15.9	906	1	CAD2_MOUSE
12	686.5	15.9	906	1	CAD2_BOVIN
13	684.5	15.8	906	1	CAD2_RAT
14	674.5	15.6	916	1	CAD4_HUMAN
15	671.5	15.5	902	1	DSC3_MOUSE
16	667.5	15.4	901	1	DSC2_HUMAN
17	667.5	15.4	906	1	CADG_XENLA
18	661	15.3	896	1	DSC3_BOVIN
19	657	15.2	883	1	CADH_HUMAN
20	653.5	15.1	863	1	DSC2_BOVIN
21	647.5	15.0	905	1	CADN_XENLA
22	647	15.0	714	1	CADN_MOUSE
23	642	14.9	893	1	DSC3_MOUSE
24	634	14.7	887	1	CAD1_CHICK
25	630.5	14.6	712	1	CAD1_CHICK
26	609	14.1	886	1	CAD1_RAT
27	608	14.1	790	1	CADI_HUMAN
28	605	14.0	4530	1	FATH_HUMAN
29	604	14.0	884	1	CADI_MOUSE
30	601.5	13.9	3579	1	STAN_DRONE
31	597.5	13.8	732	1	CADL_CHICK
32	595.5	13.8	886	1	DSCI_MOUSE
33					

34	591	13.7	882	1	CADI_HUMAN	P12830 homo sapien
35	590	13.7	872	1	CADI_XENLA	P30944 xenopus lae
36	587	13.6	784	1	CADP_MOUSE	P33146 mus musculu
37	585.5	13.6	893	1	DSCI_BOVIN	O01107 bos taurus
38	580.5	13.4	880	1	CADP_XENLA	P33148 xenopus lae
39	576	13.3	3014	1	CLRI_HUMAN	P09466 homo sapien
40	574.5	13.3	796	1	DSC2_MOUSE	P55288 mus musculu
41	574.5	13.3	894	1	DSCI_HUMAN	O08554 homo sapien
42	571.5	13.2	789	1	CADA_CHICK	P79995 gallus gall
43	571	13.2	3034	1	CLRI_MOUSE	O35161 mus musculu
44	567.5	13.1	788	1	CADI_HUMAN	O94688 homo sapien
45	565	13.1	814	1	CADP_HUMAN	P55291 homo sapien

ALIGNMENTS

RESULT 1
CADH_HUMAN STANDARD; PRT; 832 AA.
AC 012864; 015336;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cadherin-17 precursor (Liver-intestine-cadherin) (LI-cadherin)
DE (Intestinal peptide-associated transporter HPT-1).
GN CDH17.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC TISSUE-Colon adenocarcinoma;
RX MEDLINE=94204643; PubMed=8153632;
RA Dantzig A.H., Hoskins J., Tabas L.B., Bright S., Shepard R.L.,
RA Jenkins I.L., Duckworth D.C., Sportsman J.R., Mackensen D.,
RA Rostack P.R. Jr., Skatrud P.L.;
RT "Association of intestinal peptide transport with a protein related to
RT the cadherin superfamily";
RL Science 264:430-433 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Boettlinger A., Kreft B., Fieger C., Dlouhy B., Berndorf D.,
RA Goessens R., Tauber P.;
RT "Molecular cloning of human LI-cadherin: evidence for a novel type
RT of cadherin within the cadherin superfamily";
RL Submitted (DEC-1994) to the EMBL/GenBank/DBD databases.
CC -!- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
CC MANNER IN CONNECTING CELLS. CADHERINS MAY THIS CONTRIBUTE TO THE
CC SORTING OF HETEROGENEOUS CELL TYPES. LI-CADHERIN MAY HAVE A ROLE
CC IN THE MORPHOLOGICAL ORGANIZATION OF LIVER AND INTESTINE.
CC INVOLVED IN INTESTINAL PEPTIDE TRANSPORT.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE GASTROINTESTINAL TRACT AND
CC PANCREATIC DUCT. NOT DETECTED IN KIDNEY, LUNG, LIVER, BRAIN,
CC ADRENAL GLAND AND SKIN.
CC -!- SIMILARITY: CONTAINS 7 CADHERIN DOMAINS.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U07969; AAA19021.1; -;
CC EMBL: X83228; CAA58231.1; -;
CC HSSP: P15116; INCO.
CC DR GENEW; HGNC:1756; CDH17.
CC MIM: 603017; -;

InterPro: IPR002126; Cadherin.
 DR Pfam: PF000028; Cadherin; 7.
 DR PRINTS: PR00205; CADHERIN.
 DR SMART: SM00112; CA; 5.
 DR PROSITE: PS00232; CADHERIN 1; 3.
 DR PROSITE: PS50268; CADHERIN 2; 6.
 KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KM Signal; Transmembrane.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 832 CADHERIN-17.
 FT DOMAIN 23 787 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 788 808 POTENTIAL.
 FT DOMAIN 809 832 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 129 128 CADHERIN 1.
 FT DOMAIN 129 244 CADHERIN 2.
 FT DOMAIN 245 340 CADHERIN 3.
 FT DOMAIN 341 449 CADHERIN 4.
 FT DOMAIN 450 566 CADHERIN 5.
 FT DOMAIN 567 667 CADHERIN 6.
 FT DOMAIN 668 777 CADHERIN 7.
 FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 546 546 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 587 587 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 722 722 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 115 115 E -> K (IN REF. 2).
 FT CONFLICT 446 446 I -> T (IN REF. 2).
 FT CONFLICT 734 734 D -> E (IN REF. 2).
 FT CONFLICT 739 739 A -> E (IN REF. 2).
 SQ SEQUENCE 832 AA; 92147 MW; C6DB6DE0E8BA1277 C6C64;
 Query Match 100.0%; Score 4321; DB 1; Length 832;
 Best Local Similarity 100.0%; Pred. No. 7.1e-256;
 Matches 832; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

541 FGKYNASSFAKFTLLVTVNBPSPQSHVFOAKVSEDAVIGTKGVNTAKDEGLDISY 600
 541 FGKYNASSFAKFTLLVTVNBPSPQSHVFOAKVSEDAVIGTKGVNTAKDEGLDISY 600
 601 SLRGDTRGWLKIDHVTGEIFSVAPLDREAGSPRYOVVATEVGGSSLSVSEFHLIMDV 660
 601 SLRGDTRGWLKIDHVTGEIFSVAPLDREAGSPRYOVVATEVGGSSLSVSEFHLIMDV 660
 661 NDNPRLAKYTGIFPCHPLSAGSLIFKATDDOHLFRGPHFTFSLGSSGLNDMEVSK 720
 661 NDNPRLAKYTGIFPCHPLSAGSLIFKATDDOHLFRGPHFTFSLGSSGLNDMEVSK 720
 721 INGTARLSTRHTDPEERAVVVLIRINDGGRPLRGLVSLPTFCSCVEGSCFRPAHQOT 780
 721 INGTARLSTRHTDPEERAVVVLIRINDGGRPLRGLVSLPTFCSCVEGSCFRPAHQOT 780
 781 GIPTVGAVGILLTTLVIGIILAVVFRIRKDKGKDNVSAQASEVKPLRS 832
 781 GIPTVGAVGILLTTLVIGIILAVVFRIRKDKGKDNVSAQASEVKPLRS 832

RESULT 2
 CADD_MOUSE STANDARD; PRT; 827 AA.
 ID CADD_MOUSE
 AC 09R100;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cadherin-17 precursor (Liver-intestine-cadherin) (LI-cadherin) (BILL-cadherin) (P130).
 GN CDH17.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Swiss Webster; TISSUE=Intestine;
 RA Angres B., Kim L., Tauber R.;
 RT "LI-cadherin gene expression during intestinal development.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.; SEQUENCE OF 26-33; 52-58; 74-81; 117-123 AND
 RP 490-509, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RC STRAIN=BALB/c; TISSUE=Retal liver;
 RX MEDLINE=20469471; PubMed=10906147;
 RA Ohnishi K., Shimizu T., Kanaeayama H., Melchers F.;
 RT "The identification of a nonclassical cadherin expressed during B cell
 RT development and its interaction with surrogate light chain.";
 RL J. Biol. Chem. 275:31134-31144(2000).
 RL J. Biol. Chem. 275:31134-31144(2000).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES. LI-CADHERIN MAY HAVE A ROLE
 CC IN THE MORPHOLOGICAL ORGANIZATION OF LIVER AND INTESTINE.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (potential).
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IS FOUND IN INTESTINE WITH
 CC LOWER EXPRESSION IN SPLEEN, BONE MARROW, LUNG AND TESTIS. NO
 CC EXPRESSION DETECTED IN LIVER, KIDNEY, HEART, BRAIN OR SKELETAL
 CC MUSCLE. EXPRESSED IN PRECURSOR B-CELLS AND MYELOID CELLS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION INCREASES IN PRO- AND PRE-B-1
 CC CELLS, DECREASES IN LARGE AND SMALL PRE-B-11 CELLS, AND INCREASES
 CC AGAIN IN IMMATURE AND MATURE B-CELLS.
 CC -1- SIMILARITY: CONTAINS 7 CADHERIN DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 DR EMBL; AF177669; AAD51125.1; -
 DR EMBL; D87912; BAB03264.1; -
 DR HSSP; P15116; INCI.
 DR MGD; MG1:1095414; Cdh17.
 DR InterPro; IPR002126; Cadherin.
 DR Pfam; PF00028; cadherin; 7.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 5.
 DR PROSITE; PS00232; CADHERIN_1; 3.
 DR PROSITE; PS50268; CADHERIN_2; 5.
 DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KW Signal; Transport.
 FT SIGNAL 1 25
 FT CHAIN 26 827
 FT DOMAIN 26 786
 FT TRANSMEM 787 807
 FT DOMAIN 808 827
 FT DOMAIN 29 127
 FT DOMAIN 128 243
 FT DOMAIN 244 339
 FT DOMAIN 340 448
 FT DOMAIN 449 565
 FT DOMAIN 566 666
 FT DOMAIN 667 776
 FT CARBOHYD 148 148
 FT CARBOHYD 249 249
 FT CARBOHYD 418 418
 FT CARBOHYD 545 545
 FT CARBOHYD 573 573
 FT CARBOHYD 586 586
 FT CARBOHYD 721 721
 SQ SEQUENCE 827 AA; 91645 MW; CDCECEASA7652B58 CRC64;

Query Match 79.8%; Score 3450; DB 1; Length 827;
 Best Local Similarity 79.2%; Pred. No. 8.5e-203;
 Matches 658; Conservative 75; Mismatches 94; Indels 4; Gaps 1;

2 ILQAHLSLCLMLYATAGGEGKFGKPLKPMFTSYEGOEPSQIFQKAMPATPE 61
 1 MVAHQHFLCLTLTLTCGIGBEKSGPLKPMFTSIFBEOEPSQVIFQKMPATPE 60
 62 LTGETDNIFVIEBGLLYNRALDRETRSTHNLQVALANGILVEGVPVITTEVKDIND 121
 61 LTGETDGIKIEKDGILYHTRALDRETRAVHHQLALDSHGAIVGSPVITTEVKDIND 120
 122 NRPTFLOSKKEGVRONSRGKPFLLYNATLDDPATNGQLYYQIVTQLPMINWYFQ 181
 121 NRPTFLOSKKEGVRONSRGKPFMYNATLDDPATNGQLFYQIVTQLPQINDWYFQ 180
 182 INKRTGALSLTRGSGOELNPAKPSYVLVSVDKMGSGSNSPDTTSVDIITENIWK 241
 181 IDKRTGALSLTRGSGOELNPAKPSYVLVSVDKMGSGSNSPDTTSVDIITENIWK 240
 242 PKEVENENSTDPHPKITQVRWMDPGAQYSLVDKELPRFPSPSIDQEGDIYVTOPLDRE 301
 241 PEVEIRENSTDPHPKITQVQWMDPGAQYSLVNDKELSPFPSPSIDQEGDIYVTOPLDRE 300
 302 EKQAYFYAVAKBEGKPLSYPLFIHYKQVINDNPTCPSPYTVFEVQENRKLNSIGT 361
 301 EKNSHFFPAKDKENGKPLAYPEIYKVIDINDNPTCTSPYTVFEVQENRKLNSIGT 360
 362 LTAHDEDEENTANSFLNRYIVEQTPKLPMOGLFLIQTVYGMQLAQSLLKQOTPOVNT 421
 361 FEAHDEDEENANSILKTKLVDDTPKPSDGLFLIGERYEKVLSQSLKQSPQYNLS 420
 422 IEVSDKDFKTLCEVQINVIDINDQIPIFEKSDYGNLTAEADTIGSTIITQATDADEPF 481
 421 IEVSDVDFKTLCEVQINVIDINDQIPIFEKSNYGSNTLSADTIGSTIITQATDADEPF 480
 482 TGSSTKLHILIKDSEGRGLGVTDPTHTNGYVILIKPLDEFAVSNYFKAEPEPLVF 541
 481 TGSSTKLHILIKDSEGRGLGVTDPTHTNGYVILIKPLDEFAVSNYFKAEPEPLVF 540

QY 542 GVKKNASPAKFLIYTDVNEAPFOSHOVFOAKVSEDAIGTKVNTAKDPGLDYSYS 601
 DB 541 GIEYNASSPASFELIYTDVNEVFPORIPQANVSEDAVAGSVANTADPGLTYSYS 600
 QY 602 LRGDTRGMLKIDVTGHEIFSVAPLDRBAGPYRVQVAVTEVGGSSLSVSEPHILMDVN 661
 DB 601 LKGNMRGMLKIDVTGHEIFSAAPLDRBESVYVQVAVTEVGGSSLSSTADPFLVLTVN 660
 QY 662 DNPRLAKDYTGIFCHPLSAPSSLFEATDDQHLFRGHFPPSLGSGSLQNDWEYSKI 721
 DB 661 DNPRLAKDYTGIFCHPLSAPSSLFEVYDDQSLRRKFTFALGREGLQSDWEYSKI 720
 QY 722 NGTHARLSTHTPEERAAVVLFRINDGGRPLEGIYSLPVTFSCVBSGCFRPAHQNG 781
 DB 721 NGTHARLSTHTPEERAAVVLFRINDGGRPLEGIYSLPVTFSCVBSGCFRPAHQNG 780
 QY 782 IPTVGNVAVGILLTTLTVIGIILAVFIRIKDKGKNVESAQAQSEVPLRS 832
 DB 781 IPTVGNVAVGILLTTLTVIGIILAVFIRIKDKGKNVESAQAQSEVPLRS 827

RESULT 3
 CADH_RAT STANDARD; PRT; 827 AA.
 AC P55281;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cadherin-17 precursor (Liver-intestine-cadherin) (LI-cadherin).
 GN CDH17.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=9426966; PubMed=8207063;
 RA Berndorff D., Gessner R., Kreft B., Schnoy N., Lajous-Better A.-M.,
 RA Loch N., Reutter W., Hortsch M., Tauber R.;
 RT "Liver-intestine cadherin: molecular cloning and characterization of
 RT a novel Ca(2+)-dependent cell adhesion molecule expressed in liver
 RT and intestine";
 RL J. Cell Biol. 125:1353-1369(1994).
 CC - FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES. LI-CADHERIN MAY HAVE A ROLE
 CC IN THE MORPHOLOGICAL ORGANIZATION OF LIVER AND INTESTINE.
 CC - SUBCELLULAR LOCATION: Type I membrane protein.
 CC - TISSUE SPECIFICITY: Liver and intestine.
 CC - SIMILARITY: CONTAINS 7 CADHERIN DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X78997; CAA5631.1; -
 DR HSSP; P15116; INCI.
 DR InterPro; IPR002126; Cadherin.
 DR Pfam; PF00028; cadherin; 7.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 6.
 DR PROSITE; PS00232; CADHERIN_1; 3.
 DR PROSITE; PS50268; CADHERIN_2; 6.
 KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KW Signal.
 FT SIGNAL 1 21
 FT SIGNAL. POTENTIAL.

```

FT CHAIN 22 827 CADERIN-17.
FT DOMAIN 22 786 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 787 807 POTENTIAL.
FT DOMAIN 808 827 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 29 127 CADERIN 1.
FT DOMAIN 128 243 CADERIN 2.
FT DOMAIN 244 339 CADERIN 3.
FT DOMAIN 340 448 CADERIN 4.
FT DOMAIN 449 565 CADERIN 5.
FT DOMAIN 566 666 CADERIN 6.
FT DOMAIN 667 776 CADERIN 7.
FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 249 249 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 721 721 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 827 AA; 91862 MW; E1A0B03646562C01 CRC64;

Query Match 79.0%; Score 3414; DB 1; Length 827;
Best Local Similarity 78.8%; Pred. No. 1,3e-200;
Matches 655; Conservative 70; Mismatches 102; Indels 4; Gaps 1;

QY 2 ILQAHSLCLMLLYATGYGQEGKFGSPLEKMTFSTYEGQSPSOITFOFANPAPATFE 61
DB 1 MWSAQHFLCLLYTGANGQEGKFGSPLEKMTFSTYEGQSPSOITFOFANPAPATFE 60
QY 62 LTGETDNI FVIEREGLLYNNALDRSTRNLQVALDANGIIVEGVPITIEVKIND 121
DB 61 LTGETDGI FPKIEKGLLYHTRVLDRETRAVAHLLQALMDSGALVDGPVPIIEVKIND 120
QY 122 NRPFLOSKEGSRONSRPEKFLYVATLDLDPATPNQGLYQIYIQLPMINWYFQ 181
DB 121 NRPFLOSKEGSRONSRPEKFLYVATLDLDPATPNQGLYQIYIQLPMINWYFQ 180
QY 182 INNTGASLTRRESOELNPAKNPSYMLVISVKGMGQSGSFSDTTSVDITVENIMKA 241
DB 181 INNTGASLTRRESOELNPAKNPSYMLVISVKGMGQSGSFSDTTSVDITVENIMKA 240
QY 242 PKPEVWENSTDPHPIKITQVRANDPGAQVSLVDKEKLPREFSIDEGDIYVQPLDRE 301
DB 241 PEPEIHEHNLDPPIKITQVOMNDPGAHYSLINKERKLPQPFISIDEGNIYVQPLDRE 300
QY 302 EKDAVYVYAAKDEYKPLSLPLSIHYKVKDINDNPPICSPVYVFEQENRKLGNISGT 361
DB 301 EKSHVFPATKXDRNGKPLAYPLRIYKVIDINDNPPICSLQVYVFEQENRKLGNISGT 360
QY 362 LTAHDBRENTANGFLNRYIYEOTPKLPMGLFLIOTYAGMLQAKOSLKKQDTPQYNLT 421
DB 361 PAHDMDBANNINSLIKRLVDQTPKPSDBLFLIDEGYKQGLKSKQDSPQYNLT 420
QY 422 IEVSDKPKTLCPVOINVINDINDOIPFEKSDYGNLTAEADTNGISTILLIQTADDEPF 481
DB 421 IEVSDKPKTLCPVOINVINDINDOIPFEKSDYGNLTAEADTNGISTILLIQTADDEPF 480
QY 482 TGSSEKILYHITKGSSEGLGVTDPHNTGYVILKKPLDEFTAAVSNVFEAKNPEPLVYF 541
DB 481 TGSSEKILYHITKGSSEGLGVTDPHNTGYVILKKPLDEFTAAVSNVFEAKNPEPLVYF 540
QY 542 GVKYNASFAKFTLVTDVNEAPQSOHVFOAKSEDEVALSTKYGNAVTAPOEGADISYS 601
DB 541 GIEYNASSFASFTLVTDVNEAPQSOHVFOAKSEDEVALSTKYGNAVTAPOEGADISYS 600
QY 602 LRGDTRGMLKIDHYGEIFSVAPLDRBAGSPYQVAVATEVGGSSLSVSEFHLILMDVN 661
DB 601 LRGDTRGMLKIDHYGEIFSVAPLDRBAGSPYQVAVATEVGGSSLSVSEFHLILMDVN 660
QY 662 DNPRLAKDYGLGFPCPLSPAGSLIFPATDDOGLFFGPHFTSLSGSGSLQNDWEYSKI 721
DB 661 DNPRLAKDYGLGFPCPLSPAGSLIFPATDDOGLFFGPHFTSLSGSGSLQNDWEYSKI 720
QY 722 NGTHARLSTRHTDEERAYVVLIRINDGRPLBSIVSLPVTFGSCVSGSCFRPAGHQTG 781

```

```

DB 721 NGTHARLSTRHTDEERAYVVLIRINDGRPLBSIVSLPVTFGSCVSGSCFRPAGHQVG 780
QY 782 IFTVGNAGVILLTTLTVIGIILAVFIRIKDKGNVESAQSEYKPLRS 832
DB 781 IFTVGNAGVILLTTLTVIGIILAVFIRIKDKGNVESAQSEYKPLRS 827

RESULT 4
CADD RABBIT STANDARD; PRT; 829 AA.
ID CADD RABBIT STANDARD; PRT; 829 AA.
AC Q28634;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Caderin-16 precursor (Kidney-specific caderin) (Ksp-caderin).
GN CDH16.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN (1)
RP SEQUENCE FROM N.A., AND OF SEQUENCE 570-586.
RC STRAIN=New Zealand white;
RX MEDLINE=95340560; Pubmed=7615566;
RA Thomson R.B., Igarashi P., Biemesderfer D., Kim R., Abu-Alfa A.,
RA Soleimani M., Atkinson P.S.;
RT "Isolation and cDNA cloning of Ksp-caderin, a novel kidney-specific
RT member of the caderin multigene family.";
RL J. Biol. Chem. 270:17594-17601(1995).
CC -1- FUNCTION: CADERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
CC -1- THEY PREPREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
CC MANNER IN CONNECTING CELLS. CADERINS MAY THIS CONTRIBUTE TO THE
CC SORTING OF HETEROGENEOUS CELL TYPES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: KIDNEY-SPECIFIC. LIMITED TO THE BASOLATERAL
CC MEMBRANES OF RENAL TUBULAR EPITHELIAL CELLS.
CC -1- SIMILARITY: CONTAINS 6 CADERIN DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U28945; AAC48472.1; -.
DR HSBP; P09803; ISUH.
DR InterPro; IPR002126; Caderin.
DR Pfam; PF00028; caderin; 6.
DR SMART; SM00112; CA; 6.
DR PROSITE; PS00232; CADERIN_1; 2.
DR PROSITE; PS50268; CADERIN_2; 6.
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal.
FT SIGNAL 1 18
FT CHAIN 19 829
FT DOMAIN 19 786 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 787 807 POTENTIAL.
FT DOMAIN 808 829 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 126 CADERIN 1.
FT DOMAIN 131 235 CADERIN 2.
FT DOMAIN 242 336 CADERIN 3.
FT DOMAIN 341 449 CADERIN 4.
FT DOMAIN 445 564 CADERIN 5.
FT DOMAIN 569 665 CADERIN 6.
FT DOMAIN 666 786 CADERIN 7.
FT CARBOHYD 517 517 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 709 709 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 722 722 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 829 AA; 88827 MW; D2DF106C47A43B9 CRC64;

```


QY 297 PLDBEADAVFVAVAKDEYKPLSYPLEIHVVKYKDINDNPPCPSPVTVPEVOENRLG 356
 DB 293 ELDBEADAEYLOVRAQNSHGEDYAPLELHVLMWMDNDVPIPPDPVTVPELSPG 352
 QY 357 NSIGTLTAHNRDEBENTANSFLNRYIVEQTPCLPMDG-LFLIOTYAGMLQAKOSLKR-OD 414
 DB 353 TEVTRLSAEDADAPGNSHVVYVOLLSPEDGEGEAFQVDPRTSGSVTLGVLPLRAGQ 412
 QY 415 TPQYNLTLEVSXKD-KFTLCFQVQINVIDINDQIPTEKSDYGNLTLAEDTNIGSTLTI 472
 DB 413 ILLIVLMDLAGAEGFSSICEVAVATDINDHAPETTSQIGPISLPEDVEPGLVAML 472
 QY 473 QARDAD-EPTGSSKIL-YHIKGDSSGRGVDDPHTNNGYVLIK-KPLDEFTAVSN 528
 DB 473 TALDADLEP--ARLMDPALERDGTGTGLDMEP--DSGHVRLCKLKNLSYBAASHE 527
 QY 529 IVFKAENPEPLFVGVKNYASSFAFETLLIVTDVNEAPQPSQHVPOAKYSEDAIGTKVGNV 588
 DB 528 VVVVVGVAKTV-GPGGPGATAVTVLVERVMPPKLDQDSYASVPISAPNSFLLTI 586
 QY 589 TAKDPGLDISYSLRGDTGKWKIDHTGEIFSVAPLD-REAGSPRYQVATEVGGSSL 647
 DB 587 QPSDPISRTLRFSLVNDSGWLCEKESGEVHTAQSLQAGPQDPTVTVLEAODTDEPRL 646
 QY 648 SSVSEFHLIMDVNDNPRKAKOYTGLEFCHPLAPSLIF-EATDDQHLFRGPHFTFS 706
 DB 647 SASAPLVYHFLKAPPAALTLAPVPSQYLCRPHDHLIVSGPKDPLASGHGP-XSFT 705
 QY 707 LG-SGSLQNDWENSKINGTHARLSTRHTDPEERAVVLLIRINDGRRPLEGIVSLPYTFC 765
 DB 706 LGRPPTYQORDWRLOTLNGSHAVYLLALHWEPREHIIIPVVSHNAQ--WMQLLVRYVC 762
 QY 766 SC-VEGSCFPRAGHQTGIPYGMVGLITLLVYIGIILAVFRIKKDKKNVESQA 824
 DB 763 RCNVEGQCMRKVGMKMPKLSAVGLVGLVAGIIFLLIFTHMTSRKKDPOPADS 822
 QY 825 SEVK 828
 DB 823 VPLK 826

RESULT 6
 CADG_MOUSE
 ID CADG_MOUSE STANDARD; PRT; 830 AA.
 AC 088338; Q9JL25;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cadherin-16 precursor (Kidney-specific cadherin) (Ksp-cadherin).
 GN Cdh16.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RX MEDLINE=98389630; PubMed=9721215;
 RA Thomson R.B., Ward D.C., Quaggin S.B., Igarashi P., Muckler Z.E.,
 RT "cDNA cloning and chromosomal localization of the human and mouse
 RT isoforms of ksp-cadherin."
 RL Genomics 51:445-451(1998).
 RP [2]
 RC SEQUENCE OF 1-25 FROM N.A.
 RC STRAIN=129/Sy;
 RA Whyte D.A., Thomson R.B., Nix S.L., Zaniani R., Li C., Karp S.L.,
 RA Aronson P.S., Igarashi P.,
 RT "Ksp-cadherin gene promoter. I. Sequence analysis and renal
 RT epithelial-cell-specific activity."
 RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.

CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
 CC -1- TISSUE SPECIFICITY: KIDNEY-SPECIFIC.
 CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF016271; AAC34254.1; -
 DR EMBL; AF118228; AAF28836.1; -
 DR HSSP; P09803; 1SUH.
 DR MGI; MGI:106671; Cdh16.
 DR InterPro; IPR002126; Cadherin.
 DR Pfam; PF00028; cadherin; 6.
 DR SMART; SM00112; CA; 6.
 DR PROSITE; PS00232; CADHERIN_1; 1.
 DR PROSITE; PS00268; CADHERIN_2; 5.
 KM Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KW Signal.
 FT SIGNAL. 1 21
 FT CHAIN 22 830
 FT DOMAIN 22 788
 FT TRANSSEM 789 809
 FT DOMAIN 810 830
 FT DOMAIN 27 128
 FT DOMAIN 133 237
 FT DOMAIN 244 338
 FT DOMAIN 343 451
 FT DOMAIN 457 566
 FT DOMAIN 571 667
 FT DOMAIN 668 788
 FT DOMAIN 519 519
 FT CARBOHYD 604 604
 FT CARBOHYD 724 724
 SQ SEQUENCE 830 AA; 89859 MW; 5EB1F06B5B7C2F25 CRC64;
 Query Match 20.3%; Score 877.5; DB 1; Length 830;
 Best Local Similarity 30.5%; Pred. No. 5, 1e-46;
 Matches 245; Conservative 141; Mismatches 353; Indels 63; Gaps 22;
 QY 62 LTGETD---NIFVIERE-GLLYNRALDRETSHTNLQVVALDANGIIVEGPVITIEV 116
 DB 59 LSGDSNADQNTFAVPTDSEFLVATRTLDRBEKAEYQLQVTLSEDRILMGQVLTVHV 118
 QY 117 KDINDNPRFTLQSKTSGSVFQNSRPGKPRFYVNAATLDDPATNGQLYVIVQLPAINN 176
 DB 119 KDENDQVQPSQAIYRAQLSQGTRPGVPLFLASDGDAGTANSDLRFHILSQSPQPL 178
 QY 177 VMYFOINNKGAISLRBGSQELNPAKNPSYNVIVISYKMGQSGSENSFDTTSVDIIVTE 236
 DB 179 PMFQLDPRHIGALALSSGSTSLDHALSEYQLQVYKMGDPSGHQALAT-VEISIVE 237
 QY 237 NIKKAPRQVEMVNSTDPHPKIKITQVRMNDPGAQYSLVDEKLPPEFISIDEGDIYVQ 296
 DB 238 NSMAPLEPVHIANLKVYYPHISAQVHWSGDVHYQL--ESQPPGFVDTEGMLHVTM 294
 QY 297 PLDBEADAVFVAVAKDEYKPLSYPLEIHVVKYKDINDNPPCPSPVTVPEVOENRLG 356
 DB 295 ELDBEADAEYLOVRAQNSHGEDYAPLELHVLMWMDNDVPIPPDPVTVPELSPG 354
 QY 357 NSIGTLTAHNRDEBENTANSFLNRYIVEQTPCLPMDG-LFLIOTYAGMLQAKOSLKR-OD 409
 DB 355 TEVTRLSAEDADAPGNSHVVYVOLLSPEDGEGEAFQVDPRTSGSVTLGVLPLRAGQ 414
 QY 410 LKQDTPQYNLTLEV--SDRQFTLCFQVQINVIDINDQIPTEKSDYGNLTLAEDTNIGS 467

Db 415 ILIQ-----VLAVDLASBEGSLSTCEVTWMTDNNHABEFINSIGIPVTLPEDEVKGA 469
 Qy 468 TILTIQATDAD-BEFTGSSKIL-YHIKDSSEGLGVDTDPHTNTGYIILKKPLDEFTAA 525
 Db 470 LVATLTATDADLEP---AFRLMOPALIEGDEBEGFDLSWEPDSHVOGLRLKULSTYAAP 526
 Qy 526 VSNIVFAENPEBPLVFGVKNASSFAKFTLLIVDVNAFOPSOHQVAKVSEDAVIGTKV 565
 Db 527 DHKVVVVVSNIEELV-GPGCPAPATATVTLIVERVAVPLKIDQSSYETSIPVSTPAGSL 565
 Qy 586 GNTAPDREGLDISYLRGDRGLKIDHTGTGELFSAFAPLD-REAGSPRYQVYA--TEV 642
 Db 586 LTIQPSDPMRSRTIRFSLVNDSEGLCIKEVSGEHTAQSLQAGAPGDTYVTLVEAODTK 645
 Qy 643 GSSSLSSVSEFHLI-----LMDVNDNPPRLAKDYGLFCHPLSAGSLIFEATDDQH 696
 Db 646 PGLSTATVIVIHPLKASPVPAITLSAGPSR-----HICTRQDYGVAVSGVSEPD 697
 Qy 697 LFRGPHFTSLG-SGLSDNWEVSKINGTARLSTRHTDEERAYVLIIRI-NDGGRPL 754
 Db 698 ANRNGPYSFALGPMPVQDRMLQPLNDSHAVLTLALHWEPGEYVAVVHHDTMMQL 757
 Qy 755 EGVLSLPYTRCSC-VESSCFRPAHQGTGPTVGAAGVILLTLLVIGILLAVPIRL--- 810
 Db 758 Q---VAVIVCRNVEGQCMKXGMRKGMPTKLSAVGVLLGLTAAIGFILVETHLALA 813
 Qy 811 KKDQKDNVESAQSEVKPLRS 832
 Db 814 RKDL-----QPADSVPLKA 828
 RESULT 7
 CAD4 CHICK STANDARD; PRT; 913 AA.
 AC P245Q3;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cadherin-4 precursor (Retinal-cadherin) (R-cadherin) (R-CAD).
 GN CDH4.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 ON NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Retina;
 RX MEDLINE=91299341; PubMed=1712604;
 RA Inuzuka H., Miyatani S., Takeichi M.;
 RT "R-cadherin: a novel Ca(2+)-dependent cell-cell adhesion molecule
 RT expressed in the retina.";
 RT Neuron 7:69-79(1991).
 RL
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES. MAY PLAY AN IMPORTANT ROLE
 CC IN RETINAL DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: EMBRYONIC BRAIN AND NEURONAL RETINA.
 CC -1- DEVELOPMENTAL STAGE: DETECTED ONLY AFTER SOME DEGREE OF NEURONAL
 CC DIFFERENTIATION HAS TAKEN PLACE AND PERSISTS AT LEAST UP TO THE
 CC NEWLY HATCHED STAGE.
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D14459; BAA03356.1; -.
 DR PIR; JH0424; ICHCR.
 DR HSSP; P15116; INCI.
 DR InterPro; IPR002126; Cadherin.
 DR InterPro; IPR000233; Cadherin_C_term.
 DR Pfam; PF00028; cadherin; 5.
 DR Pfam; PF01049; Cadherin_C_term; 1.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 5.
 DR PROSITE; PS00232; CADHERIN 1; 3.
 DR PROSITE; PS0268; CADHERIN 2; 5.
 DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KW Signal.
 FT SIGNAL 1 19
 FT PROPEP 20 166
 FT CHAIN 167 913
 FT DOMAIN 167 731
 FT TRANSMEM 732 753
 FT DOMAIN 754 913
 FT DOMAIN 167 274
 FT DOMAIN 275 389
 FT DOMAIN 390 504
 FT DOMAIN 505 610
 FT DOMAIN 611 721
 FT DOMAIN 721 885
 FT CARBOHYD 280 280
 FT CARBOHYD 409 409
 FT CARBOHYD 554 554
 FT CARBOHYD 629 629
 FT CARBOHYD 658 658
 FT CARBOHYD 699 699
 FT VARIANT 652 652
 SQ SEQUENCE 913 AA; 100885 MW; BD2BB9EC0815DD6 CRC64;
 Query Match 16.5%; Score 715; DB 1; Length 913;
 Best Local Similarity 26.9%; Pred. No. 4.7e-36;
 Matches 203; Conservative 117; Mismatches 329; Indels 106; Gaps 19;

Qy 123 RPTPLQSKTEGSGVRSRGRKPELVY-----NATDLDPATPNQGLYYQIY 169
 Db 31 KGFSEEDYTAFAFSQVIMEGQKLKVFNNCAKNGVRYETNSLDKVRADGTMY--AVH 88
 Qy 170 QLPMINWVYFOI-----NKGTAISLTRGSOB 198
 Db 89 QVQMASKQILLMTWMTDPTGLRWEALVRFVGEKIQHNGHKPKGKSGPVLLAQOOSDT 148
 Qy 199 LNPAPKPSYNLVISVKMGQSGSENSFSDTSVDIITENIMKAPKEVEMVENSDDHPK 258
 Db 149 LLPWRQHQ-----SAKGLRQKRD-----WVLP-PINVPENSRGPPQ 186
 Qy 259 ITQVRNN---DGAQVSL--VDEKELRPFPISIDQ-EGDIYTOPLDREKAYVYVYA 312
 Db 187 LVRIKSDKKEIHIRISITGVAGADQPMFVFISIDPVGMYVTRPDRREERASYHLRA 246
 Qy 313 KQVEGKPLSYPLIEHVKNVDINDNPTCPSPVTFVEQENELKNSIGTLTHADREENT 372
 Db 247 VDMNGKVENPDPIDVYIVDAMDNRPEFNGVYNGSVDEGSGRTGYVMTVMTANDADSTT 306
 Qy 373 ANSFLNRIVEQTPKLPMDGLFLIQTYAGMLQALQSLKKQDTPQNLTIKESDKD---- 428
 Db 307 ANGMVYRIVTQDPSQSNMFTINSETDITVAAGLDREKQGMVIVQATDMEGNIN 366
 Qy 429 --FKTLCFQVQVINDNDPIPEKSDYGNULIADTNIGSTIITLIQADDADEPFGSSK 486
 Db 367 YGLSTNTALITVTDVNDNPPFTTSTSG--EVENRVEVVANLTVDNRQPSHPNN 424
 Qy 487 ILVHIKDSSEGLGVDTDPHTNTGYIILKKPLDEFTAAVSNIVFAENPEBPLVFGVKYN 546
 Db 425 AIVRIISGSPSHFTIRIDPVTNNEGMYVYKAVDTEMRAFLTLTWVNSQALASGIGWS 484
 Qy 547 ASGPAFTLLIVDVNEAFOQSVFOAKVSEDAVIGTKVGNVYAKDPEGL---DISYSLR 603
 Db 485 POSTAGVTISVDVNEAEPFPTNHLIRLEEGVPIGTVLTTPSAVDPRFMQOAVRYSKL 544

QY 604 GDRGWLKIDHVTGEISVAPLDREA-----GSPRVQVATEVGGSSLSVSEHLLMD 659
 DB 545 SDAWNNININATNGQITTAVALDDESDYIKNNVTEATFELADNGIPASGTGLQITLID 604
 QY 660 VDNPPRLADYDGLFCFCHPLSAPGSLIFEAT-----DDOHLFRGPHFTSLGS--GSLQ 713
 DB 605 INDNAPLLEKRAOI-----CEKNLAVINITADADIDPRV--GP-FVELEBPSPSAVR 656
 QY 714 NDWEVSKINGTHARLSTRHDFEERAVVLIIRINDGRPPLEGIVSLPTFCSCVSGSCF 773
 DB 657 KNNWITRLNGDYAQLSRIMYLEAGYVDPLIVTDSGNPPLYNTSIKVKPCDENGDC 716
 QY 774 RPAGHOTGIPVGAAGVGLTLTLVIGIILAVPI 808
 DB 717 TTIG--AVAAAGTGAIIAILICITLITLTVLL 748

RESULT 8
 ID CAD4_MOUSE STANDARD; PRT; 913 AA.
 AC P39038;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cadherin-4 precursor (Retinal-cadherin) (R-cadherin) (R-CAD).
 GN C0H4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreatic islets;
 RX MEDLINE=94067164; PubMed=8247017;
 RA Hutton J.C., Christofori G., Chi W.X., Edman U., Guest P.C.,
 Hanahan D., Kelly R.B.;
 RT "Molecular cloning of mouse pancreatic islet R-cadherin: differential
 RT expression in endocrine and exocrine tissue.";
 RL Mol. Endocrinol. 7:1151-1160 (1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94095672; PubMed=8270638;
 RA Matsunami H., Miyatani S., Inoue T., Copeland N.G., Gilbert D.,
 Jenkins N.A., Takeichi M.;
 RT "Cell binding specificity of mouse R-cadherin and chromosomal mapping
 RT of the gene.";
 RL J. Cell Sci. 106:401-409 (1993).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES. MAY PLAY AN IMPORTANT ROLE
 CC IN RETINAL DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: DISTRIBUTED WIDELY IN MOUSE TISSUES WITH HIGH
 CC LEVELS PRESENT IN BRAIN, SKELETAL MUSCLE, AND THYMUS.
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X69966; CAA49589.1; -
 DR EMBL; D14888; BAA03605.1; -
 DR PIR; A47543; A47543.
 DR HSSP; P15116; INCI.
 DR MGD; MGI:99218; Cdh4.
 DR InterPro; IPR002126; Cadherin.
 DR InterPro; IPR000233; Cadherin_C_term.

DR Pfam; PF00028; cadherin; 5.
 DR Pfam; PF01049; Cadherin_C_term; 1.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 5.
 DR PROSITE; PS00232; CADHERIN_1; 3.
 DR PROSITE; PS0268; CADHERIN_2; 5.
 DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KW Signal.
 FT SIGNAL. 1 20
 FT PROPEP 21 166
 FT CHAIN 167 913
 FT DOMAIN 167 731
 FT TRANSMEM 732 753
 FT DOMAIN 754 913
 FT DOMAIN 167 274
 FT DOMAIN 275 389
 FT DOMAIN 390 504
 FT DOMAIN 505 610
 FT DOMAIN 611 721
 FT DOMAIN 721 885
 FT CARBOHYD 146 146
 FT CARBOHYD 280 280
 FT CARBOHYD 409 409
 FT CARBOHYD 554 554
 FT CARBOHYD 629 629
 FT CARBOHYD 658 658
 FT CARBOHYD 699 699
 SQ SEQUENCE 913 AA; 100030 MW; 1245A8CE8C38481 CRC64;
 Query Match 16.3%; Score 705; DB 1; Length 913;
 Best local similarity 30.1%; Pred. No. 1.9e-35;
 Matches 178; Conservative 101; Mismatches 280; Indels 33; Gaps 12;
 QY 239 WKAPKEVMEVNSTDPHPIKITQVRW---NDPAQVSL--VDKELPRFESID-QEGDI 232
 DB 168 WVLP-PINVENSRGPFPOOLVIRSDKDDIPRYSITGVGADQPMVEVNDISGSM 226
 QY 293 YVTPQPLDREKDYVFAVAKDEYKPLSYLEIHYVKVDINDNPFCSPVTVFEVOEN 352
 DB 227 YVTRPMDREERASVHLRAHVDANGKVENPIDLYIVIDMNNRBEFINQVNGSDEG 286
 QY 353 ERLGNSIGTLTAHDEENTANSFLNRYIEQTPKLPMDGLPLIQYAGMLQAKSLRK 412
 DB 287 SKPGTYVMTYANDADSTTANGMVRVRYVTPQSPSQMFTINSETGDIIVVAAGLDR 346
 QY 413 QDTPQVNLTEVSDK-----FKTLFVQINVIDINDQPIREKSDYGLTLAEDNIG 466
 DB 347 EKVQOQVIVQADTMEGNLNYGLSNTATALLTVTVANDNPEFTSTFAG--EVENRIE 404
 QY 467 STLTQATADAEPTGSSKILYHIKGDSEGRGVDTDPHTNGVYLIKKPLDFETAAV 526
 DB 405 TVVANLTVMRDQPHSNMNAVRIISGDSGFVSRTDVTNMEGNTVAKADYELNRA 464
 QY 527 SNIVFAKNENPELPVFGVKNASSPAKFTLLVTVDNEAPQFSQHVQAKVSEDAVIGTVG 586
 DB 465 FMLTVMWSNQAPLASGIQMSFGSTAGVTISVTVNNEAPYPSNHKILRLREGVAGTALT 524
 QY 587 NVTAKDEGL---DISLSLGDTRGWLKIDHVTGEISVAPLDREA-----GSPRVQVVA 639
 DB 525 TFSADVDDRMQOAVRYSKISDPANMHAITSNGQITTAALDRESLYTKNNVTEATFEL 584
 QY 640 TEVGGSSLSVSEHLLMDVNDPRLADYDGLFCFCHPLSAPG--SLIFEATDDOHL 697
 DB 585 ADNQPISPAAGTGLQITLIDINDAPLLEKRAOI-----CERPLNAINITADADMDP 639
 QY 698 FRGPH-FTFSLGSGSLQNDWEVSKINGTHARLSTRHDFEERAVVLIIRINDGRPPLEG 756
 DB 640 NIGYVVELFPIPTYAKMTITRLNGDYAQLSRIMYLEAGYVDVPIIVTDSGNPPLSN 699
 QY 757 IVSLPVTFCSCVSGSCRPAGHOTGIPVGAAGVGLTLTLVIGIILAVPI 808
 DB 700 TSVIKVAVCPDENGDCCTTVG---AVAAAGTGAIIAILICITLITLTVLL 748

DB 494 NPEFAPNPKIRQESGLHAGTMLTFTADPDRYQONIRYTKLSDPANMLKIDPVNGOI 553
 QY 620 FSVAPIDREA---GSPRYQVAVATVGGSSLSVSEFHLIMDVNDVNPRLAKVYGLF 675
 DB 554 TTIYAVIDRESPPVKNINATFLASDNGIIPMSGTCTLOIYLDIDNAPDVL----- 606
 QY 676 FCHPLSA-----PGSLIFEATDDQHLFRGPH-FTFSIGSGSLONDEVSKINGTAR 727
 DB 607 ---POBAECETPDPSINTALDYIDPAGFPADLPSPPTIKRNTITRLNDGFAQ 663
 QY 728 LSTRHTDFERAVVILIRINDGRPPLEGIVSLPTVFCSC-VEGSCFRAGHOTGIPV- 785
 DB 664 LNLKIKLEAGIYEVEPIITITDSGNPPKSNISIRVYQCQDSNGDC-----TDVDRIV 716
 QY 786 ---GMANGILLTTLVGIILAVFIRIKKDKGDNVSAQASEK 829
 DB 717 GAGLGTGAILIALLIITLILVLMFVWMKRRDKERAKOLLIDP 762

RESULT 10
 CAD2_CHICK STANDARD; PRT: 912 AA.
 AC P10288; Q90630;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neural-cadherin precursor (N-cadherin) (Cadherin-2).
 GN CMH2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NC NCBI_TaxID=9031;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8153917; PubMed=2833236;
 RA Hata K., Nose A., Nagafuchi A., Takeichi M.;
 RT "Cloning and expression of cDNA encoding a neutral calcium-dependent
 RT cell adhesion molecule: its identity in the cadherin gene family.";
 RL J. Cell Biol. 106:873-881(1988).
 RN (2)
 RP SEQUENCE OF 1-25 FROM N.A.
 RC STRAIN=Cornish white rock Cockerel;
 RX MEDLINE=97354288; PubMed=9210582;
 RA Li B., Paradies N.E., Brackenbury R.W.;
 RT "Isolation and characterization of the promoter region of the chicken
 RT N-cadherin gene.";
 RL Gene 191:7-13(1997).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN
 CC NEURONAL RECOGNITION MECHANISM.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X07277; CAA30258.1; -
 DR EMBL: U15563; AAB62980.1; -
 DR PIR: A29964; IJCHCN.
 DR HSRP: P15116; INCI.
 DR InterPro: IPR002126; Cadherin.
 DR InterPro: IPR000233; Cadherin_C_term.
 DR Pfam: PF00028; cadherin_5.
 DR Pfam: PF01049; cadherin_C_term; 1.
 DR PRINTS: PRO0205; CADHERIN.

DR SMART, SMO0112; CA; 5.
 DR PROSITE; PS00232; CADHERIN_1; 3.
 DR PROSITE; PS0268; CADHERIN_2; 5.
 KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KM Signal.
 FT 1 28
 FT PROPE 29 164
 FT CHAIN 165 912
 FT DOMAIN 165 729
 FT TRANSMEM 730 752
 FT DOMAIN 753 912
 FT DOMAIN 165 272
 FT DOMAIN 273 387
 FT DOMAIN 388 502
 FT DOMAIN 503 609
 FT DOMAIN 610 720
 FT DOMAIN 728 884
 FT CARBOHYD 869 884
 FT CARBOHYD 278 278
 FT CARBOHYD 330 330
 FT CARBOHYD 407 407
 FT CARBOHYD 578 578
 FT CARBOHYD 628 628
 FT CARBOHYD 657 657
 FT CONFLICT 21 21
 SQ SEQUENCE 912 AA; 100464 MW; 9BA5AC9DC1FCC489 CRC64;
 A -> G (IN REF. 2).
 Query Match 16.1%; Score 696; DB 1; Length 912;
 Best Local Similarity 26.9%; Pred. No. 6,7e-35;
 Matches 209; Conservative 117; Mismatches 328; Indels 124; Gaps 22;
 QY 126 FLQSKRGSVRQNRGRGKFLVYNATDLPATPNQGLYYQIVYQLPMINWYPOIN-- 183
 DB 41 FPEVDHSAVSRKSVHGGQPLNVRFSQDSE-----NRKIYFSSSE 81
 QY 184 -----NKTGALISLTREGSOLNPAKNPSYNLVSYKDMGGS-----ENSTSDT 228
 DB 82 EDRVEDDGVVNERGFQLSAEPT-----FVVSADKKEQEBWQKVLTPPAFTGAS 136
 QY 229 SVDIIVTENT-----WKAKPEVMEVNSTDPHPKITQVRYMNDPGA 269
 DB 137 EKQKKIEDIIFPQWQKDSHLKQKRDVIF-PINLPNSGPPQDELVRIR----- 189
 QY 270 QYSLVDNKEKLRFPFSI-----DQ-----EGDIYVQPLREERKAVVFAVAK 313
 DB 190 -----SRDKSLSRYSVTGPGADQPPGIFITINPISQSLVTKRLDREQLASHLRAHAV 245
 QY 314 DEYKPLSYPLEIHVVKVDINDNPPCTCPSPVTFEVEGNERLGNISIGTLTAHPRDENTA 373
 DB 246 DVNNGOVENIDIVINVIDNDRPEFLHGVNMGTVPEGSKPGTYVWTVAIDADDPNAQ 305
 QY 374 NSFLNRYIVHQPKLPMDGLFLQITVAGMLQAKQSLKKQDTQVYNLTIEVSKD----- 428
 DB 306 NGMLRYRLISQAQSPSPNNFTINNETGDIITVAAGIDREKVOQYTLIIQATDMEGNPTY 365
 QY 429 -FKTLGVQINVIDINDQIPIF-EKSDYGMILTAEDINIGSTIIITQATDADEPFTSSK 486
 DB 366 GLNSTATVATVTDVNDVNDNPEFTAMTYGEV---PENRVYIVANLVTDKXQDHTPAMN 422
 QY 487 ILVHIKDESEGRGLVDTPHTNTGYVILKKPLDFTFAVSNIVFAKNEPPLVGYKYN 546
 DB 423 ARYQMTGDDPRTGQITLIDNSNDGLTVVVKPIDFTNRMFVLVAENQVPLAKGIQHP 482
 QY 547 ASSPAKFTLLIVDVNEAPQSSQHVQAKVSEDAVIGKVGNVYAKDP-----EGGDISYSL 602
 DB 483 PQSTATVSIIVIDVNESPYFPNPKLVRQEBGLAGSMLTFTFARDDRYQQTSLRSK 542
 QY 603 RGDTRGMCLKIDHVTGELFSVAPIDREA-----GSPRYQVAVATVGGSSLSVSEFHLIM 658
 DB 543 LSDPANMLKIDPVNGOITTTAVLDRESIYVQNNNTATFLASDNGIIPMSGTGLQIYLL 602
 QY 659 DVNNPPRL-AKVDTGLFFCHPLSAPGSLIFEATDDQHLFRGPHFTFSLGSQ--SLQND 715
 DB 603 DINDVAPVNFKEAT--TCETLQ-PNAINTAVDPDIDPNAAG-FAFELDPSPSISRN 657

Qy	716	WEVSKNKGTHALSTHTTPEEAAVYLKININGGPPLEIGVSLPTFCSC-VEGSCFR	774
Db	658	WTVIRISGDHQAOLNIRFLKAGIVDPPIYITDSGNPHASTSVLKVKVQCQDINGDC--	715
Qy	775	PAGHOTGIFTV--GNAVGLLTLLVIGIIIAVFIPIKDKXKONVESAOSEVPR	829
Db	716	-----TDVDRIVAGAGTGTAINIILICIIILITVLVFMFVWMRRRDXEROAKOLLIDP	768
RESULT 11			
CAD2_MOUSE		STANDARD;	PRT; 906 AA.
AC	P15116;	O64260;	
DT	01-APR-1990	(Rel. 14, Created)	
DT	01-APR-1990	(Rel. 14, Last sequence update)	
DT	15-OCT-2001	(Rel. 40, Last annotation update)	
DE	Neural-cadherin precursor (N-cadherin) (Cadherin-2).		
GN	CDH2.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]	SEQUENCE FROM N.A.	
RP	MEDLINE=89346748; PubMed=2762814;		
RA	Miyetani S., Shimamura K., Hatta M., Nagafuchi A., Nose A.,		
RA	Matsumura M., Hatta K., Takeichi M.;		
RL	"Neural cadherin: role in selective cell-cell adhesion.";		
RL	Science 245:631-635(1989).		
RN	[2]	SEQUENCE FROM N.A.	
RP	Tamura K.;		
RL	Submitted (NOV-1997) to the EMBL/Genbank/DBD databases.		
RN	[3]	PARTIAL SEQUENCE FROM N.A.	
RP	STRAIN=C57BL/6;		
RC	MEDLINE=92409532; PubMed=1528849;		
RX	Miyetani S., Copeland N.G., Gilbert D.J., Jenkins N.A., Takeichi M.;		
RT	"Genomic structure and chromosomal mapping of the mouse N-cadherin		
RT	gene.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 89:8443-8447(1992).		
RN	[4]	DEVELOPMENTAL STAGE.	
RP	STRAIN=C57BL/6; TISSUE=Testis;		
RC	MEDLINE=97033837; PubMed=8879495;		
RX	Munro S.B., Blaschuk O.W.;		
RA	"A comprehensive survey of the cadherins expressed in the testes of		
RT	fetal, immature, and adult mice utilizing the polymerase chain		
RT	reaction.";		
RL	Biol. Reprod. 55:822-827(1996).		
RN	[5]	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 160-267.	
RP	MEDLINE=95191680; PubMed=7885471;		
RC	Shapiro L., Fannon A.M., Kwong P.D., Thompson A., Lehmann M.S.,		
RA	Handrickson W.A.;		
RA	Griegel G., Legrand J.-F., Als-Nielsen J., Colman D.R.;		
RT	"Structural basis of cell-cell adhesion by cadherins.";		
RT	Nature 374:327-337(1995).		
RN	[6]	X-RAY CRYSTALLOGRAPHY (3.4 ANGSTROMS) OF 160-374.	
RP	MEDLINE=96318235; PubMed=9655503;		
RA	Tamura K., Shan W.S., Hendrickson W.A., Colman D.R., Shapiro L.;		
RT	"Structure-function analysis of cell adhesion by neural (N-)		
RT	cadherin.";		
RL	Neuron 20:1153-1163(1998).		
CC	-1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.		
CC	THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC		
CC	MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE		
CC	SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN		
CC	NEURONAL RECOGNITION MECHANISM.		
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.		
CC	-1- DEVELOPMENTAL STAGE: EXPRESSED AT ALL STAGES OF TESTICULAR		

```
CC DEVELOPMENT WITH HIGHEST LEVELS FOUND IN TESTS OF 21-DAY-OLD MICE.
```

```
-I- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
```

```
-----
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
```

```
-----
```

```
DR EMBL; M31131; AAA37353.1; -.  
DR EMBL; AB0086811; BAA23549.1; -.  
DR EMBL; S45011; AAB23356.1; -.  
DR PIR; A32759; IJMSCN.  
DR PDB; INCG; 1O-JUL-95.  
DR PDB; INCJ; 1O-JUL-95.  
DR PDB; INCI; 1O-JUL-95.  
DR PDB; INCU; 18-MAR-99.  
MCD; MGI:88355; Cdh2.  
DR InterPro; IPRO02126; Cadherin.  
DR InterPro; IPRO00233; Cadherin_C_term.  
DR Pfam; PF00028; cadherin; 5.  
DR Pfam; PF01049; Cadherin_C_term; 1.  
DR PRINTS; PR00205; CADHERIN.  
DR SMART; SMO0112; CA; 5.  
DR PROSITE; PS00232; CADHERIN_1; 3.  
DR PROSITE; PS02068; CADHERIN_2; 5.  
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
KW Signal; 3D-structure.  
FT SIGNAL 1 23  
FT PROPEP 24 159 POTENTIAL.  
FT CHAIN 160 906 NEUTRAL-CADHERIN.  
FT DOMAIN 160 724 EXTRACELLULAR (POTENTIAL).  
FT TRANSSEM 725 746 POTENTIAL.  
FT DOMAIN 747 906 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 160 267 CADHERIN 1.  
FT DOMAIN 268 382 CADHERIN 2.  
FT DOMAIN 383 497 CADHERIN 3.  
FT DOMAIN 498 603 CADHERIN 4.  
FT DOMAIN 604 717 CADHERIN 5.  
FT DOMAIN 863 878 SER-RICH.  
FT CARBOHYD 190 190 N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD 273 273 N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD 325 325 N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD 402 402 N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD 572 572 N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD 651 651 N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD 692 692 N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CONFLICT 7 9 GRG -> APR (IN REF. 2).  
FT CONFLICT 565 565 Y -> N (IN REF. 2).  
FT CONFLICT 567 567 O -> K (IN REF. 2).  
FT CONFLICT 624 624 A -> T (IN REF. 2).  
SQ SEQUENCE 906 AA; 99761 MW; 0206741B77107B7 CRC64;
```

```
Query Match 16.1%; Score 694.5; DB 1; Length 906;  
Best Local Similarity 27.6%; Pred. No. 8.2e-35;  
Matches 211; Conservative 122; Mismatches 332; Indels 99; Gaps 23;
```

```
Oy 117 KDINDNPTFLQSGRTGSVFRONSRRPGKPELYVNATLDIDDPATNGQLYQIVQLPMINN 176  
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::  
Db 47 KDVEEGCP-LLNVFENCSNCRK-----RKVOYESESEPADFUKDEGVTV--AVRSFPLTAE 98  
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::  
Oy 177 VWFVQI---NNKG-----AISLTRSSGSLNPAPKNP-STNLVSVSDM---GGQSENPF 224  
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::  
Db 99 QAKFLYLIAQKEIEQEKQVAVNLISRELTLPTEPKPEHIEIIEIVFPRLAKHGSGALOROK 158  
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::  
Oy 225 SDTTSVDIIITENIWKAPEKVENWENTSDPHPIKITQVRMN-----DPGAQYS 272  
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::  
Db 159 RD-----TWIP-PINLPENSRRGPFGELVRIIRDNRDKNSLRYSVTGPGA--- 202  
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::  
Oy 273 LVDEKLEPRPPFSIDQ-EGDIYVTOPLDREEDKAYVFAVAKDYGKPLSYPLEIHVKV 331
```



```

Db      203 ----DPPPTGIFITINPISGQSLSTKPLDRELIRFHLRAHVDINGNOVENPIDIYINYI 258
Qy      332 DINDNPPTCPSPVTVPEVOENERLGNISIGTLTAHDREENTANSFLNRYIVEQTPKLPMD 391
Db      259 DMDNDNRPEFLHQWNSVPEGSRGTYVMTVTAIDADDPAALGMRLRYRLISQAFSTPSP 318
Qy      392 GFLITQYAGMLQAKOSLKKQDTPQYNLTIEVSDK-----FKLCEVQIVINDINQ 445
Db      319 NMFTINNETGDIITVAAGLDREKVOQYTLTIQATDMEGNPTGLSNATATAVITVTVNDN 378
Qy      446 IPIF-EKSDYGNLTAEADTNGSTILTIQATDDEPTGSSKILYHIKDGSEGRIGVPT 504
Db      379 PPEFTAMTFYGEV---PENRVDIVANLITVDKQPHTPMNAAYRISGSDPTGRFALIT 435
Qy      505 DPHNTGYVILIKKPLDFTAAVSNIVFKAENPEPLVFGVKYKNSAFKFTLITVDNEAP 564
Db      436 DPNNDGLTVVVKPIBETNRMFVLVAAGNOVPLAKGIQHPPOSTATVSVFVIDVNEP 495
Qy      565 QESQHVFOAKVSEDAIGTVGNVTAADPEGL--DISYSLRGDTGMLKIDHVTGEIFS 621
Db      496 YFAPNPKIRQESGLHAGTWTLTITADDPDRYMOQNRITKLSDBPAMMLKIDEPVNGQIT 555
Qy      622 VAPLIDREA---GSPYRVOVATEVGGSSLSVSEFHLIMDVNDNPPRLAKDYTGIFFC 677
Db      556 IAVLDRESPYVQNNIYATLASDNGILPMSGCTLOIYILINDNAPQVL-----606
Qy      678 HPLSA-----PGLIFBATDDOHLFRGPH-FTFSLGSGSLQNDWEVSKINGTHARLS 729
Db      607 -PGEAETCEPPEPNSINIALDYIDNAGPFAFDLPLSPVTLIKMWTINRLNDDFAQLN 665
Qy      730 TRHTDPEERAYVVLIRINDGRRPPELIVSLPTFGCG-VEGSGFRAGHGTGTPV---785
Db      666 LKIKFLKAGIYEVIITITDSGNPKNSISILRYKVCQCDNSNGDC-----TVDRLVGA 718
Qy      786 GMAVGLITLTLVIGLITAVFIRIKKQKQNDVESAQASBVR 829
Db      719 GLGTGATITALLCIIILITLIVMFVVMKRDKEROKOLLIDP 762

```

RESULT 12

```

CAD2_BOVIN STANDARD; PRT; 877 AA.
ID AC P19534;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neural-cadherin precursor (N-cadherin) (Cadherin-2) (Fragment).
GN CDH2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=30360979; PubMed=2390969;
RA Llaw C.W., Cannon C., Power M.D., Kiboneka P.K., Rubin L.L.;
RT "Identification and cloning of two species of cadherins in bovine
RT endothelial cells."
RL EMBL J. 9:2701-2708 (1990).
CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
CC SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN
CC SURFACIAL RECOGNITION MECHANISM.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
CC -----
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X53615; CAA37677.1; -.
CC DR PIR; S11693; IUBOCHN.
CC DR HSSP; P5116; INCI.
CC DR InterPro; IPR002126; Cadherin.
CC DR InterPro; IPR000233; Cadherin_C-term.
CC DR Pfam; PF00028; cadherin; 5.
CC DR Pfam; PF01049; cadherin_C-term; 1.
CC DR SMART; SM00112; CA; 5.
CC DR PROSITE; PS00232; CADHERIN_1; 3.
CC DR PROSITE; PS00268; CADHERIN_2; 5.
CC KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat.
CC
CC FT PROPEP 1 130
CC FT CHAIN 131 877
CC FT DOMAIN 131 695
CC FT TRANSMEM 696 717
CC FT DOMAIN 718 877
CC FT DOMAIN 131 238
CC FT DOMAIN 239 363
CC FT DOMAIN 364 469
CC FT DOMAIN 469 574
CC FT DOMAIN 575 685
CC FT DOMAIN 834 849
CC FT CARBOHYD 161 161
CC FT CARBOHYD 244 244
CC FT CARBOHYD 296 296
CC FT CARBOHYD 373 373
CC FT CARBOHYD 543 543
CC FT CARBOHYD 622 622
CC FT CARBOHYD 663 663
CC SQ SEQUENCE 877 AA; 96845 MW; 4418829ED871A249 CRC64;
CC
CC Query Match 15.9%; Score 688.5; DB 1; Length 877;
CC Best Local Similarity 26.7%; Pred. No. 1,8e-34;
CC Matches 203; Conservative 126; Mismatches 343; Indels 87; Gaps 19;
CC
CC Qy 126 FLOSKYEGSVRONSRBPKFLVYNATDLDPATPNQGLYYQIVQLP-----MINN 176
CC Db 7 FPEEDVYSAVLSDRVLEGGPLLVNKKFENC-----NKKRAYQESSBPADFVDEDMVYA 60
CC
CC Qy 177 VMTFQINNKTKGALSTRBESQELNPAKNSYLVISKMGSGSESPDITSVDIIV-- 234
CC Db 61 VRSFPLSSHSKFLIYAQDK-----TQERKQVAVKLSPALPEDSVKESREIEIIVP 115
CC
CC Qy 235 -----TENIWKAPKPEVWENSTDPPIKITOVNRMN-----DPCAQ 270
CC Db 116 RQVTKANGYLQROKRWVLP-PINLPENSRGPPQGLVAIRSDRDNLRLRYSVTGPGL- 173
CC
CC Qy 271 YSLVDREKLPPEPFSIDQ-EGDIYVTPQPLDREKQAVYVAVAKDEYKPLSYPLSIHYK 329
CC Db 174 -----DQPEPTGIFINPISGQSLSTKPLDRELIRFHLRAHVDINGNOVENPIDIYIN 227
CC
CC Qy 330 VKQINNPPTCPSPVTVPEVOENERLGNISIGTLTAHDREENTANSFLNRYIVEQTPKLP 389
CC Db 228 VIDMNDNRPEFLHQWNSVPEGSRGTYVMTVTAIDADDPAALGMRLRYRLISQAFSTP 287
CC
CC Qy 390 MDGLFLITQYAGMLQAKOSLKKQDTPQYNLTIEVSDK-----FKLCEVQIVINDIN 443
CC Db 288 SPNMFITINNETGDIITVAAGLDREKVOQYTLTIQATDMEGNPTGLSNATATAVITVTVND 347
CC
CC Qy 444 DQIPIF-EKSDYGNLTAEADTNGSTILTIQATDDEPTGSSKILYHIKDGSEGRIGVPT 502
CC Db 348 DNPPEFTAMTFYGEV---PENRVDIVANLITVDKQPHTPMNAAYRISGSDPTGRFALIT 404
CC
CC Qy 503 DTPHTNTGYVILIKKPLDFTAAVSNIVFKAENPEPLVFGVKYKNSAFKFTLITVDNEAP 562
CC Db 405 QTDPNNDGLTVVVKPIBETNRMFVLVAAGNOVPLAKGIQHPPOSTATVSVFVIDVNEP 464

```

QY 563 APOPSQVHFOAKSSEDAVIGTKVGNVAKDEPGI---DISYSLGDTGRGMIKIDHTVGEI 619
 DB 465 NPFAPAPPKIIRBEGHAGVLTFTFAADPDRYMOQINIRYTKSDPANMLKIDSVNGOI 524
 QY 620 FSVAPIIDREA-----GSPYRQVNVATEVGGSSLSVSEPHILIMVNNPRLAKDYGLF 675
 DB 525 TTIAVLDRSEBPNVKNATYNATFLASDNGIPMSGTGLQIYLLDINNAPOVPEQAEI 583
 QY 676 FCHPLSAPGSLIFERATDDOHLFRGPRH-FTFSLGSGSLNDWSEVSKINGTHARLSTRHTD 734
 DB 584 -CE-TPDPNSINITALDYIDDPNAGPFAFDLPSPVTKIKNTITRANGDPFAQLNKLKIF 641
 QY 735 FEEBAYVVLIRINDGPRPLEGIVSLPVTQSC-VEGSCRRPAGHOTGIPTV---GNAV 790
 DB 642 LEAGIYEVPLIITDSGNPKSNISILRKVCQCDSNGDC-----TDVDRIVAGLGTG 694
 QY 791 ILTLTLVIGIILAVFIRIKOKGKNVESAQSEKFP 829
 DB 695 AIIATILCITILILVLMFVMMKRRKXEROAQLLIDP 733

RESULT 13
 CAD2_RAT STANDARD; PRT; 906 AA.
 ID CAD2_RAT STANDARD; PRT; 906 AA.
 AC Q921Y3; Q9R0T5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neural-cadherin precursor (N-cadherin) (Cadherin-2).
 GN CMH2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
 RC STRAIN=Sprague-Dawley; TISSUE=Testis;
 RX MEDLINE=98187820; PubMed=9528971;
 RA Chung S.S., Mo M.Y., Silverstein B., Lee W.M., Cheng C.Y.;
 RT "Asi testicular N-cadherin: its complementary deoxyribonucleic acid
 cloning and regulation.";
 RL Endocrinology 139:1853-1862 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RA Asai K., Tada T., Yamamoto M., Tada A., Mizuno M., Rimoto T., Kato T.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CADHERINS ARE CALCITUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN
 CC NEURONAL RECOGNITION MECHANISM.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).
 CC -1- TISSUE SPECIFICITY: IN TESTIS, EXPRESSED IN SEROLI AND GERM
 CC CELLS.
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF097593; AAC8318.1; -;
 DR EMBL; AB017695; BAA84919.1; -;
 DR HSSP; P15116; INCU.
 DR InterPro; IPR002126; Cadherin.
 DR InterPro; IPR000233; Cadherin_C term.
 DR Pfam; PF00008; cadherin_5.
 DR Pfam; PF01049; cadherin_C term; 1.
 DR PRINTS; PR00205; CADHERIN.

DR SMART; SM00112; CA; 5.
 DR PROSITE; PS00232; CADHERIN 1; 3.
 DR PROSITE; PS0268; CADHERIN 2; 5.
 KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KM Signal.
 FT 1 27
 FT 28 159
 FT PROPEP 906
 FT CHAIN 160 724
 FT DOMAIN 160 724
 FT TRANSMEM 725 745
 FT DOMAIN 746 906
 FT DOMAIN 160 267
 FT DOMAIN 268 382
 FT DOMAIN 383 497
 FT DOMAIN 498 603
 FT DOMAIN 604 717
 FT DOMAIN 863 878
 FT CARBOHYD 190 190
 FT CARBOHYD 273 273
 FT CARBOHYD 325 325
 FT CARBOHYD 402 402
 FT CARBOHYD 572 572
 FT CARBOHYD 622 622
 FT CARBOHYD 651 651
 FT CARBOHYD 692 692
 FT CONFLICT 7 48
 FT CONFLICT 48 48
 FT CONFLICT 153 153
 FT CONFLICT 646 646
 FT CONFLICT 658 658
 FT CONFLICT 724 724
 SQ SEQUENCE 906 AA; 99685 MW; 97D9937ED8D1F8B5 CRC64;

Query Match 15.9%; Score 686.5; DB 1; Length 906;
 Best Local Similarity 27.2%; Pred. No. 2.5e-34;
 Matches 215; Conservative 126; Mismatches 344; Indels 105; Gaps 24;

QY 97 AALDANGIIV-----EGVPITIEVKDINDRPTFLQSGKGVYRNSRPGKFLYVNA 150
 DB 21 ASLHSGELALCKTGEPREYSAVLPKTYHEQP-LLVNKFNSCRNK-----RKQYESS 74
 QY 151 TDLDDPATNGQLYYQIVIQIOLPMINNVWFQI---NNKTG-----AISLRGSGOELNPA 202
 DB 75 EPADFKVDEDTIVY-AVNSPFLSABQAKFLVLAQDKXQEKQVAVNLSLPSLUEEPM 132
 QY 203 KNP-STNLVIVSKDM---GGQSENSFSDTTSVITITENIKAPKPEVENVSTDPHPYK 258
 DB 133 KSPHEIEEIVFPROLAKHSGALQROKRD-----WVTP-PINLPENSRGFPPOE 179
 QY 259 ITOVRNN-----DPAQVSLVDKELPPFPSSIDQ-EGDIYVTOPLDREKKA 305
 DB 180 LVRLNSDRKMLSLKSVTGPQA-----DPPGIGIFINPISQLSTYKPLDRELLAR 232
 QY 306 YVFVAVADKYGKPLSYPLEIHKVYKINDNBPCTSPVTVVEVOENERLGSIGTLTAH 365
 DB 233 FHLRAVAVDINQVBNPIDIYINVIDNDNRPEFLHQVWNSVSGSKPGYVMTVIAI 292
 QY 366 DDDERTANSPLNTKIVBQTPKLPMDGLFLITVAGMQLAKQSLKQDTPQYNLTIEVS 425
 DB 293 DADDPAVLANGMRYRILSQAPSTPSPMFTINNETGDIITVAAGDREKVOQYTLIQAT 352
 QY 426 DKD-----FKTLCEVOINVIDINDQIPF-EKSYGMLTLAEDNINISITLITQATAD 478
 DB 353 DNEGNPTGLSTATAVAVITVDVNNPPEFTMTTYGEV---PENRVDVYVAVNLVTYTDQ 409
 QY 479 EPTGSSKLLYHIINKGDSBGLVDTPHTNGYVLIKKPLDPETAASVNIYFKAENEP 538
 DB 410 QHTPAWMAAAYRISGDPGRFALITDPNSNDGLVTVKPIDFETNBMFVLLVAENQVP 469
 QY 539 LVFGKYNAASSPAKTLTYVDVNEAPQPSQVFAKVSEDAVIGTKVGNVAKDEPGI-- 596
 DB 470 IAKGIQHPQSTATVSVTVIVDVENPYPAPPKIIRBEGHAGVLTFTFAADPDRYMOQ 529

QY 597 -DISYSLRGDTRGMLKIDHTYGEIFSVAPLDRFA-----GSPYRQVAVTEVGSSLSVSS 651
DB 530 QNIRYITLSDPANMLKIDPNVNGQITTAVIDRSPNKNINATPLASDNGIPMSGTG 589
QY 652 EFHILIMDVNDNPRILAKDYTGLEFCHPLSA-----PQSLIFEATDDOHLFRGPH-F 703
DB 590 TLGIYILIDINDMNAQVL-----PQSAETCETEPNSINTALDYIDIPNAGPFAF 639
QY 704 TFSLGSSSLQNDWEVSKINGTHARLSTRHDFEERAAVVLIRINDGRPPLLEGVSLPYT 763
DB 640 DLPLSPATIKRMWITRLNDPFAQLNKIKFLEAGIYEVPIVITDSNPPKNSIILRVK 699
QY 764 FCSC-VEGSCFRPAGHOTGIPTV---GMAYGILLTTLVIGIILAVFIRIKDKGKNV 819
DB 700 VCCDSNGDC-----TDVRIYAGAGLGTGTIAILIILILVLMVVMKRDKE 752
QY 820 ESAQASEVKP 829
DB 753 RQAKQLIDP 762
RESULT 14
ID_CAD4 HUMAN STANDARD; PRT; 916 AA.
AC P55283;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cadherin-4 precursor (Retinal-cadherin) (R-cadherin) (R-CAD).
GN CDH4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95073006; PubMed=7982033;
RA Tanihara H., Sano K., Heilmann R.L., St John T., Suzuki S.;
RT "Cloning of five human cadherins clarifies characteristic features of
RT cadherin extracellular domain and provides further evidence for two
RT structurally different types of cadherin.";
RL Cell Adhes. Commun. 2:15-26(1994).
RN [2]
RP SEQUENCE OF 393-916 FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=91283540; PubMed=2059658;
RA Suzuki S., Sano K., Tanihara H.;
RT "Diversity of the cadherin family: evidence for eight new cadherins
RT in nervous tissue.";
RL Cell Regul. 2:261-270(1991).
CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
CC SORTING OF HETEROGENEOUS CELL TYPES. MAY PLAY AN IMPORTANT ROLE
CC IN RETINAL DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED MAINLY IN BRAIN BUT ALSO FOUND IN
CC OTHER TISSUES.
CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L34059; AAA35627.1; -
DR HSSP; P15116; INCI.
DR Genew; HGNC:1763; CDH4.
DR MIM; 603006; -.

DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR00233; Cadherin_C_term.
DR Pfam; PF00028; cadherin; 5.
DR Pfam; PF01049; Cadherin_C_term; 1.
DR PRINTS; SM00205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN 1; 2.
DR PROSITE; PS0268; CADHERIN 2; 5.
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW signal.
FT SIGNAL. 1 20
FT PROPEP 21 169
FT CHAIN 170 916
FT DOMAIN 170 734
FT TRANSMEM 735 756
FT DOMAIN 757 916
FT DOMAIN 170 277
FT DOMAIN 278 392
FT DOMAIN 393 507
FT DOMAIN 508 613
FT DOMAIN 614 724
FT DOMAIN 873 888
FT CARBOHYD 283 283
FT CARBOHYD 412 412
FT CARBOHYD 557 557
FT CARBOHYD 632 632
FT CARBOHYD 661 661
FT CARBOHYD 702 702
SQ SEQUENCE 916 AA; 100446 MW; E717C54A19E0C52A CRC64;
Query Match 15.8%; Score 684.5; DB 1; Length 916;
Best local similarity 29.3%; Pred. No. 3,4e-34;
Matches 176; Conservative 106; Mismatches 288; Indels 31; Gaps 12;
QY 239 WKAPKEVENSTDPHPKIITQVRW---NDPQAQVSL--VDKELPRPFSID-QEGDI 292
DB 171 WVIP-PINVPENSRGFPPOOLVIRISDKNDIPRYSITGVGADQPMVEFSLNSGRR 229
QY 293 YVTQPLDREKKAIVYVAADVEYKPLSPLEIHVKVDINDNPTCPSPVTVFVQEN 352
DB 230 YVTRPDREHASYHLRAHVDNNGKVENPIDLYIVDADNHEFINGVNCVDEG 289
QY 353 ERLGNSIGTLTAHDEENTANSFLNRYIVEQTPKLPMDGLFIQTAGMLQAKOSLKK 412
DB 290 SKRGTYVMTITANDADSTTANGMVARRYITQRPSPQNMFTINSETGDIIVVAQMDR 349
QY 413 QDTPQYNLTIVESDKD-----FKTLCFQVQIVINDIQIPIFEKSDYGNLTAEPTNG 466
DB 350 EKVQQTVAIVQATDMEGNLNYGLSNTATAIITVTVDNDPSEFTASTFAG--EVPENSV 407
QY 467 STLTITQATDADRPFGSSKILVHIKGSSEGLGVDTPHTNTGVYIIKKPLDFETAIV 526
DB 408 TVVANLTWMDRDPHPNNAVRIITSGDSGHSFVRIDPVNMEGVYVKAVDYELNRA 467
QY 527 SNIVFAENPEPVPFGKYVNASFAFTLVTDVNEAPQFSQHFQAKVSEDAVIGTKVG 586
DB 468 FMILTWVSNQAPLASIQMSFQSTAGVTISINDINAPVPSHKILRLBEGVPPTVLT 527
QY 587 NVTAKDPEGI--DISYSLRGDTRGWLKIDHTYGEIFSVAPLDRFA-----GSPYRQV 639
DB 528 TFSADVDPDRMQQAVVYSKIDPASWLNATNGQITTAVIDRSPNKNINATPLASDNGIP 587
QY 640 TFGVSSLSVSEFHLIMDVNDNPRILAKDYTGLEFCHPLSAPSLIFEATDDOHLFR 699
DB 588 ADNGIPPASGTGLQIYILIDINDNAPBLPKAQICERPLNA--INITAADADVHPNI 644
QY 700 GPH-FTFSLGSSSLQNDWEVSKINGTHARLSTRHDFEERAAVVLIRINDGRPPLLEG 758
DB 645 GPVYFELPFPYPAVRKNWITRLANGDYAQLSLILVLEAGMVVPIIVTDSGNPPLSNTS 704
QY 759 SLPTVFCSEVSGCFRPAHQGTIPVVGAAVGLTTLVIGIILAV--FIRIKDKGK 816
DB 705 IIKVVCPCDNDGDCITIG---AVAAAGLGTGAIVAIIICILITLITVLLFVVMKRDKE 761

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2003, 20:21:49 ; Search time 8328 Seconds

(without alignments)
11689.347 Million cell updates/sec

Title: US-10-025-380-1076

Perfect score: 3345

Sequence: 1 gaattccgctcgcacacatg.....cctttgcggccgcgaattc 3345

Scoring table: IDENTITY NUC
Gapop 10%0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_in:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rtd:*

36: em_htg_man:*

37: em_htg_vit:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	3345	100.0	3345 6 AX193509	AX193509 Sequence
2	3345	100.0	3345 6 AX330272	AX330272 Sequence
3	3345	100.0	3345 6 AX332252	AX332252 Sequence
4	3345	100.0	3345 6 AX410642	AX410642 Sequence
5	3345	100.0	3345 6 HSU07969	U07969 Human Intes
6	3314.8	99.1	3654 6 AX330312	AX330312 Sequence
7	3314.8	99.1	3654 6 AX335495	AX335495 Sequence
8	3314.8	99.1	3654 6 AX411163	AX411163 Sequence
9	3314.8	99.1	3654 6 HSRNMLTCA	X03228 H.sapiens m
10	2499	74.7	2499 6 140307	140307 Sequence 2
11	2499	74.7	2499 6 RNCADHER	X78997 R.norvegicu
12	1774.4	53.0	3628 10 RNCADHER	D87912 Mus musculi
13	1758.6	52.6	3414 10 D87912	AF177669 Mus muscu
14	1747.8	52.3	3351 10 R4TDPDPT	L46874 Rattus norv
15	1184.8	35.4	1733 10 R4TDPDPT	AP003478 Homo sapi
16	848.8	25.4	153894 9 AP003478	AC021576 Homo sapi
17	848.8	25.4	166450 9 AP003351	AP003351 Homo sapi
18	847.2	25.3	176854 9 AP003351	AX340772 Sequence
19	514.2	15.4	523 6 AX340772	G26817 human STS S
20	508.8	15.2	1066 11 G26817	AX341845 Sequence
21	497.2	14.9	501 6 AX341845	AX340324 Sequence
22	469.4	14.0	495 6 AX340324	AX339851 Sequence
23	461.6	13.8	492 6 AX339851	AX396194 Sequence
24	431.4	12.9	516 6 AX396194	AX339931 Sequence
25	410.4	12.3	568 6 AX396490	AX396490 Sequence
26	365.6	10.9	370 6 AX339931	AX3397080 Sequence
27	342	10.2	342 6 AX3397080	AX396419 Sequence
28	340.4	10.2	500 6 AX396419	AX260426 Sequence
29	331.8	9.9	335 6 AX260426	AX261891 Sequence
30	331.8	9.9	335 6 AX261891	AX192454 Sequence
31	331.4	9.9	346 6 AX192454	AX193263 Sequence
32	331.4	9.9	346 6 AX193263	AF428098 Danio rer
33	319.8	9.6	325 6 AF428098	AX339825 Sequence
34	306	9.1	3373 6 AX339825	AX339794 Sequence
35	269.6	8.1	274 6 AX339794	AX351289 Sequence
36	265.2	7.9	265 6 AG025648	AG025648 Homo sapi
37	265	7.9	265 6 HS12803	Z98742 Human DNA s
38	243.4	7.3	422 17 HS12803	AC008438 Homo sapi
39	211.6	6.3	119350 9 AC008438	AC0080184 Homo sapi
40	210.6	6.3	85718 9 AC0080184	AC011399 Homo sapi
41	210.6	6.3	147078 2 AC011399	AC004605 Homo sapi
42	210.6	6.3	174158 9 HUAC004605	AC009994 Homo sapi
43	210.6	6.3	259474 2 AC118269	AC118269 Homo sapi
44	210.2	6.3	229752 2 AC118269	
45	209.6	6.3	129758 9 AC118269	

ALIGNMENTS

RESULT 1

AX193509 3345 bp DNA linear PAT 15-AUG-2001

LOCUS DEFINITION Sequence 1076 from Patent WO0149716.

AX193509

ACCESSION AX193509

VERSION AX193509.1 GI:15211449

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J., Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.

AUTHORS

TITLE Compounds for immunotherapy and diagnosis of colon cancer and

Pred. No. is the number of results predicted by chance to have a

QY 1981 CCAATATGGGTACAGTGTGGCCACAGAGTAGGGGGTCTTCTTAAGCTGTGTCA 2040
 DB 1981 CCAATATGGGTACAGTGTGGCCACAGAGTAGGGGGTCTTCTTAAGCTGTGTCA 2040
 QY 2041 GAGTTCACCGTATCCCTTATGATGATGAAGCAACCTCCCGAGGTAGCCAGAGTAC 2100
 DB 2041 GAGTTCACCGTATCCCTTATGATGATGAAGCAACCTCCCGAGGTAGCCAGAGTAC 2100
 QY 2101 ACGGGCTTGTCTTCTCCATCCCTCAGTGCACCTGGAAGTCTCATTTTGAAGTACT 2160
 DB 2101 ACGGGCTTGTCTTCTCCATCCCTCAGTGCACCTGGAAGTCTCATTTTGAAGTACT 2160
 QY 2161 GATGATATACGACCTTATTTCCGGGTCCTCCATTTTCAATTTTCCCTGCGAGTGAAGC 2220
 DB 2161 GATGATATACGACCTTATTTCCGGGTCCTCCATTTTCAATTTTCCCTGCGAGTGAAGC 2220
 QY 2221 TTACAAAAGACCTGGGAAGTTTCCAAAATCAATGATCTCATGCGGCACTGTACACAG 2280
 DB 2221 TTACAAAAGACCTGGGAAGTTTCCAAAATCAATGATCTCATGCGGCACTGTACACAG 2280
 QY 2281 CACACAGACTTTGAGAGAGGCGTATGTCATTTCCGATCCATGATGAGGGGTGCG 2340
 DB 2281 CACACAGACTTTGAGAGAGGCGTATGTCATTTCCGATCCATGATGAGGGGTGCG 2340
 QY 2341 CCAACCTTGAAGGCAATGTTTCTTACAGTTACATTTGCAAGTGTGTGAAGAGT 2400
 DB 2341 CCAACCTTGAAGGCAATGTTTCTTACAGTTACATTTGCAAGTGTGTGAAGAGT 2400
 QY 2401 TGTTCCTGGCCAGAGTCAACAGCTGGATACCACTGTGGGCACTGGAGTTGTATA 2460
 DB 2401 TGTTCCTGGCCAGAGTCAACAGCTGGATACCACTGTGGGCACTGGAGTTGTATA 2460
 QY 2461 CTGCTGACCAACCTTCTGTGATGTGATTAATTTTACAGTGTGTATCCGATTAAG 2520
 DB 2461 CTGCTGACCAACCTTCTGTGATGTGATTAATTTTACAGTGTGTATCCGATTAAG 2520
 QY 2521 AAGGATTAAGGCAAGATTAATTTGAAAGTCTCAAGCATCTGAAGTCAAACTCTGAGA 2580
 DB 2521 AAGGATTAAGGCAAGATTAATTTGAAAGTCTCAAGCATCTGAAGTCAAACTCTGAGA 2580
 QY 2581 AGCTGAATTTGAAAGAAATGTTGAATTTATATAGAAATGCTATTTTCGCAACCA 2640
 DB 2581 AGCTGAATTTGAAAGAAATGTTGAATTTATATAGAAATGCTATTTTCGCAACCA 2640
 QY 2641 TCTCATCCCTTATCTTTTCACTAAGTGCATTAATTTTAAACAGATTTCCCTCT 2700
 DB 2641 TCTCATCCCTTATCTTTTCACTAAGTGCATTAATTTTAAACAGATTTCCCTCT 2700
 QY 2701 TGTCTTATATATTTGCTAAATATTTCTTTTGAAGTGAAGTCTGTCTGTGCGCAG 2760
 DB 2701 TGTCTTATATATTTGCTAAATATTTCTTTTGAAGTGAAGTCTGTCTGTGCGCAG 2760
 QY 2761 GCTGAGTACAGTGTGTATCCAGTCACTGCAACCTCCGCTCTGTGGTTCAATGA 2820
 DB 2761 GCTGAGTACAGTGTGTATCCAGTCACTGCAACCTCCGCTCTGTGGTTCAATGA 2820
 QY 2821 TTCTCCGCTCAGCTTCTCAAGTGTGCTTAAAGGCAACCAACCAATCCCAAGCT 2880
 DB 2821 TTCTCCGCTCAGCTTCTCAAGTGTGCTTAAAGGCAACCAACCAATCCCAAGCT 2880
 QY 2881 AATTTTGTATTTTATATAGACGGGGTTTCGCAATTTGGCCAGGCTGTGTGAAGT 2940
 DB 2881 AATTTTGTATTTTATATAGACGGGGTTTCGCAATTTGGCCAGGCTGTGTGAAGT 2940
 QY 2941 CTGACGTCAAGTATCTGCTGCTGTGTCTCCCAATACAGGCACTGAACACACCA 3000
 DB 2941 CTGACGTCAAGTATCTGCTGCTGTGTCTCCCAATACAGGCACTGAACACACCA 3000
 QY 3001 CCTACTTAATATTTATGATGTATAGACATTAAGAGATTTTTCATTTTCCATAGACAT 3060
 DB 3001 CCTACTTAATATTTATGATGTATAGACATTAAGAGATTTTTCATTTTCCATAGACAT 3060
 QY 3061 TTTTCTCTCTGCAAAATGAGCTTACTGTTGTTTCTTCTTGGGGCAAGACAGACT 3120

DB 3061 TTTTCTCTCTGCAAAATGAGCTTACTGTTGTTTCTTCTTGGGGCAAGACAGACT 3120
 QY 3121 CATTAATATTCGTACATTTTCTTATCAAGAGATATACAGTGTGTCTATAGA 3180
 DB 3121 CATTAATATTCGTACATTTTCTTATCAAGAGATATACAGTGTGTCTATAGA 3180
 QY 3181 ACTGCTGATTCATTTATGTTTCTTATCAATTCATTCCTGTCTCCCTCATCTTGAC 3240
 DB 3181 ACTGCTGATTCATTTATGTTTCTTATCAATTCATTCCTGTCTCCCTCATCTTGAC 3240
 QY 3241 TCTTTGTATTTTCACTGAATTTCAAACTTTGTCAAGAGAGAGAGAGAGAGTACTCAG 3300
 DB 3241 TCTTTGTATTTTCACTGAATTTCAAACTTTGTCAAGAGAGAGAGAGAGTACTCAG 3300
 QY 3301 GAAAAATTAATTAATAAGAACAGCTTTTGGCGCGGAATTC 3345
 DB 3301 GAAAAATTAATTAATAAGAACAGCTTTTGGCGCGGAATTC 3345

RESULT 2

AX330272

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Query Match 100.0%; Score 3345; DB 6; Length 3345;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGCTGACCACTGATGAGAGAAAGACTTTAAACCAATTTGTGACTTA 60
 DB 1 GAATTCGCTGACCACTGATGAGAGAAAGACTTTTAAACCAATTTGTGACTTA 60
 QY 61 CAGAAAGATTTGAATTAAGAAACTATGATATCTTCAAGCCATCTTCACTCCGTGT 120
 DB 61 CAGAAAGATTTGAATTAAGAAACTATGATATCTTCAAGCCATCTTCACTCCGTGT 120
 QY 61 CAGAAAGATTTGAATTAAGAAACTATGATATCTTCAAGCCATCTTCACTCCGTGT 120
 DB 61 CAGAAAGATTTGAATTAAGAAACTATGATATCTTCAAGCCATCTTCACTCCGTGT 120
 QY 121 CTTCTTATGCTTTATTTGGCAACTGATATGSCCAAGAGGGAGTTAGTGAACCCCTG 180
 DB 121 CTTCTTATGCTTTATTTGGCAACTGATATGSCCAAGAGGGAGTTAGTGAACCCCTG 180
 QY 121 CTTCTTATGCTTTATTTGGCAACTGATATGSCCAAGAGGGAGTTAGTGAACCCCTG 180
 DB 121 CTTCTTATGCTTTATTTGGCAACTGATATGSCCAAGAGGGAGTTAGTGAACCCCTG 180
 QY 181 AAAACCATGACATTTTCTATTTATGAGGCAAGAAACGAGTCAATTTATTCAGTTT 240
 DB 181 AAAACCATGACATTTTCTATTTATGAGGCAAGAAACGAGTCAATTTATTCAGTTT 240
 QY 181 AAAACCATGACATTTTCTATTTATGAGGCAAGAAACGAGTCAATTTATTCAGTTT 240
 DB 181 AAAACCATGACATTTTCTATTTATGAGGCAAGAAACGAGTCAATTTATTCAGTTT 240
 QY 241 AAGGCAATCTCTGCTGTGACTTTTGAATTAAGTGGGAGACAGACATATTTGTG 300
 DB 241 AAGGCAATCTCTGCTGTGACTTTTGAATTAAGTGGGAGACAGACATATTTGTG 300
 QY 241 AAGGCAATCTCTGCTGTGACTTTTGAATTAAGTGGGAGACAGACATATTTGTG 300
 DB 241 AAGGCAATCTCTGCTGTGACTTTTGAATTAAGTGGGAGACAGACATATTTGTG 300
 QY 301 ATAGAACGGGAGGACTTCTGTATTTCAACAGAGCTTTGGAAGAGAAACAGATCTACT 360
 DB 301 ATAGAACGGGAGGACTTCTGTATTTCAACAGAGCTTTGGAAGAGAAACAGATCTACT 360

QY 361 CAGATCTCAGGTTGGAGCCCTGGACGCTTAATGAATTTATAGTGGAGGGTCCAGTCCCT 420
| | | | |
Db 361 CAGATCTCAGGTTGGAGCCCTGGACGCTTAATGAATTTATAGTGGAGGGTCCAGTCCCT 420
QY 421 ATCAACATAGAAGTGAAGACATCAACGACATGACCCAGTTTCTCCAGTCMAAGTAC 480
| | | | |
Db 421 ATCAACATAGAAGTGAAGACATCAACGACATGACCCAGTTTCTCCAGTCMAAGTAC 480
QY 481 GAAGGCTCAGTAAAGCGAACTCTGCCGAGAAAGCCCTTTGTATGTCAATGCCACA 540
| | | | |
Db 481 GAAGGCTCAGTAAAGCGAACTCTGCCGAGAAAGCCCTTTGTATGTCAATGCCACA 540
QY 541 GACCTGATATCCGCGACCTCCCAATGGCAGTTTATTAACGATTTGATCCAGCTT 600
| | | | |
Db 541 GACCTGATATCCGCGACCTCCCAATGGCAGTTTATTAACGATTTGATCCAGCTT 600
QY 601 CCCATGATCAACATGTCATGTAATCTTCAATCAACAAACGAGGAGCATCTCTT 660
| | | | |
Db 601 CCCATGATCAACATGTCATGTAATCTTCAATCAACAAACGAGGAGCATCTCTT 660
QY 661 ACCCGAGAGGATCTCAGAAATGTAATCTGCTAAGAAATCTTCTTAATCTGTGATC 720
| | | | |
Db 661 ACCCGAGAGGATCTCAGAAATGTAATCTGCTAAGAAATCTTCTTAATCTGTGATC 720
QY 721 TCAGTGAAGACANTGGAGGCGAGATGAGAAATTCCTCACTGATATACCATCTGTGAT 780
| | | | |
Db 721 TCAGTGAAGACANTGGAGGCGAGATGAGAAATTCCTCACTGATATACCATCTGTGAT 780
QY 781 ATCATAGTGAACAGAAATATTTGGAAAGACAAACCTGTGAGATGTGGAAATCTCA 840
| | | | |
Db 781 ATCATAGTGAACAGAAATATTTGGAAAGACAAACCTGTGAGATGTGGAAATCTCA 840
QY 841 ACTGATCTCTCAACCCATCAAAATATCTAGGTGCGGTGGAATGATCCCGGTGCAATAT 900
| | | | |
Db 841 ACTGATCTCTCAACCCATCAAAATATCTAGGTGCGGTGGAATGATCCCGGTGCAATAT 900
QY 901 TCCCTAGTGAACAAAGAAAGCTGCGCAAGATTCCTCAATTTGCAATGACCAAGAGAGAT 960
| | | | |
Db 901 TCCCTAGTGAACAAAGAAAGCTGCGCAAGATTCCTCAATTTGCAATGACCAAGAGAGAT 960
QY 961 ATTACGTGACTGACCCCTTGGACCGAGAGAAAGATGATATGTTTTTATGACAGTT 1020
| | | | |
Db 961 ATTACGTGACTGACCCCTTGGACCGAGAGAAAGATGATATGTTTTTATGACAGTT 1020
QY 1021 GCAAAGATGATAGCGAAACCACTTCAATCTGCGTGAATTCATGTAAGTTAA 1080
| | | | |
Db 1021 GCAAAGATGATAGCGAAACCACTTCAATCTGCGTGAATTCATGTAAGTTAA 1080
QY 1081 GATATTAATGATATCAACCTACANTGTCCTGACACGATACCGTATTTGAGGTCCAGAG 1140
| | | | |
Db 1081 GATATTAATGATATCAACCTACANTGTCCTGACACGATACCGTATTTGAGGTCCAGAG 1140
QY 1141 AATGACGACTGGGTAAACAGTATCGGACCTTACTGCAATGACAGGATGAAGAAAT 1200
| | | | |
Db 1141 AATGACGACTGGGTAAACAGTATCGGACCTTACTGCAATGACAGGATGAAGAAAT 1200
QY 1201 ACTGCCAACAGTTTCTTAACCTAAGATGTGTGACAAATCTCCCAATCTGTGAT 1260
| | | | |
Db 1201 ACTGCCAACAGTTTCTTAACCTAAGATGTGTGACAAATCTCCCAATCTGTGAT 1260
QY 1261 GGAATCTTCTTAATCCAAACCTATGCTGGAATGTTCAGTTAGCTTAACAGTCTGTGAG 1320
| | | | |
Db 1261 GGAATCTTCTTAATCCAAACCTATGCTGGAATGTTCAGTTAGCTTAACAGTCTGTGAG 1320
QY 1321 AAGCAAGATCTCTCAGTACAACTTAACGATAGAGGTGTGCAAAAGATTTCAAGACC 1380
| | | | |
Db 1321 AAGCAAGATCTCTCAGTACAACTTAACGATAGAGGTGTGCAAAAGATTTCAAGACC 1380
QY 1381 CTTGTTTGTGCAAAATCAAGTTATGATATCAATGATCAATGCCATCTTTGAAAAA 1440
| | | | |
Db 1381 CTTGTTTGTGCAAAATCAAGTTATGATATCAATGATCAATGCCATCTTTGAAAAA 1440
QY 1441 TCAGATTAATGAACCTGACTCTTGCTGAGAACCAAACTTTGGGTCCACATCTTAACC 1500
| | | | |

Db 1441 TCAGATTAATGAACCTGACTCTTGCTGAGAACCAAACTTTGGGTCCACATCTTAACC 1500
QY 1501 ATCCAGGCCACTGATGCTGATGAGCCATTTACTGGAGATTTCTTAATTTCTGATCATATC 1560
| | | | |
Db 1501 ATCCAGGCCACTGATGCTGATGAGCCATTTACTGGAGATTTCTTAATTTCTGATCATATC 1560
QY 1561 ATTAAGGAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCATACCAACCGGA 1620
| | | | |
Db 1561 ATTAAGGAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCATACCAACCGGA 1620
QY 1621 TATGTCATTAATTAAGAGCTCTTGAATTTGAAACAGAGCTGTTTCCAACTGTGTC 1680
| | | | |
Db 1621 TATGTCATTAATTAAGAGCTCTTGAATTTGAAACAGAGCTGTTTCCAACTGTGTC 1680
QY 1681 AAGCAGAAATCTCTGAGCTCTAGTGTGTGTGTAAGTACATACGAGTTCTTTGCC 1740
| | | | |
Db 1681 AAGCAGAAATCTCTGAGCTCTAGTGTGTGTGTAAGTACATACGAGTTCTTTGCC 1740
QY 1741 AAGTCACGCTTAATGTAACAGATGTAAGGACCTCAATTTTCCCAACAGTATTC 1800
| | | | |
Db 1741 AAGTCACGCTTAATGTAACAGATGTAAGGACCTCAATTTTCCCAACAGTATTC 1800
QY 1801 CAAGCAAAAGTCAAGTGAAGATGCTAATAGGCACTAAAGTGGGCAATGATGCTGCAAG 1860
| | | | |
Db 1801 CAAGCAAAAGTCAAGTGAAGATGCTAATAGGCACTAAAGTGGGCAATGATGCTGCAAG 1860
QY 1861 GATCCGAAAGTCTGACATTAAGCTTTCACTGAGGGGAGACACAAAGAGTTGGCTTAA 1920
| | | | |
Db 1861 GATCCGAAAGTCTGACATTAAGCTTTCACTGAGGGGAGACACAAAGAGTTGGCTTAA 1920
QY 1921 ATTGACACGCTGATCTGATGATCTTATGATGCTCAATTTGAGAGAGAGCCGGAAGT 1980
| | | | |
Db 1921 ATTGACACGCTGATCTGATGATCTTATGATGCTCAATTTGAGAGAGAGCCGGAAGT 1980
QY 1981 CCATATCGGGTACAAAGTGTGGCCACAGAAAGTGAAGGGGGTCTTCTTAAGCTGTGTCA 2040
| | | | |
Db 1981 CCATATCGGGTACAAAGTGTGGCCACAGAAAGTGAAGGGGGTCTTCTTAAGCTGTGTCA 2040
QY 2041 GAGTTCACCTGATCTTATGATGATGATGACAAACCTCCAGGCTAGCCAGAGCTAC 2100
| | | | |
Db 2041 GAGTTCACCTGATCTTATGATGATGATGACAAACCTCCAGGCTAGCCAGAGCTAC 2100
QY 2101 ACGGGCTGTGTTCTTGCAATCCCTCAGTGCACCTGGAAGTCTCATTTTCAGGCTACT 2160
| | | | |
Db 2101 ACGGGCTGTGTTCTTGCAATCCCTCAGTGCACCTGGAAGTCTCATTTTCAGGCTACT 2160
QY 2161 GATGATGATCAGCACTTATTCGGGGTCCCATTTTACATTTTCCCTCGGCACTGGAAGC 2220
| | | | |
Db 2161 GATGATGATCAGCACTTATTCGGGGTCCCATTTTACATTTTCCCTCGGCACTGGAAGC 2220
QY 2221 TTACAAAACGACTGGGAGATTTCCAAATATCAATGATATCCCGACTGTCTACAGG 2280
| | | | |
Db 2221 TTACAAAACGACTGGGAGATTTCCAAATATCAATGATATCCCGACTGTCTACAGG 2280
QY 2281 CACACAGACTTTAAGAGAGGGGCGTATGTCCTGATCCGATCAANTATGAGGGGTGCG 2340
| | | | |
Db 2281 CACACAGACTTTAAGAGAGGGGCGTATGTCCTGATCCGATCAANTATGAGGGGTGCG 2340
QY 2341 CCAACCTTGGAGGCAATGTTTCTTAAACAGTTACATTTGCAAGTGTGTGGAAGAGT 2400
| | | | |
Db 2341 CCAACCTTGGAGGCAATGTTTCTTAAACAGTTACATTTGCAAGTGTGTGGAAGAGT 2400
QY 2401 TGTTCGCGGACAGAGGTCAACAGACTGGGATCCCACTGTGGGCAATGGAGTTGTGATA 2460
| | | | |
Db 2401 TGTTCGCGGACAGAGGTCAACAGACTGGGATCCCACTGTGGGCAATGGAGTTGTGATA 2460
QY 2461 CTGCTGACCAACCTTCTGAGATGTGATATTTTACAGTTGTGTTATCCGATTAAG 2520
| | | | |
Db 2461 CTGCTGACCAACCTTCTGAGATGTGATATTTTACAGTTGTGTTATCCGATTAAG 2520
QY 2521 AAGGATTAAGCAAGTATATGTTGAAGTCTCAAGCATCTGAAGTCAAACTCTGAGA 2580
| | | | |

```

Db 2521 AAGGATAAGGCAAGATTAATGTTGAAAAGTGCTCAAGCATCTGAATGCAAACTCTGAGA 2580
Qy 2581 AGCTGAATTTGAAAGAGATGTTGAATTTATATAGCAAGTCTATTTCAGCAACAACA 2640
Db 2581 AGCTGAATTTGAAAGAGATGTTGAATTTATATAGCAAGTCTATTTCAGCAACAACA 2640
Qy 2641 TCTCATCTTAATCTTTTCACTTCAAGTCAATTAATTTTAAACAGATATTCCTCT 2700
Db 2641 TCTCATCTTAATCTTTTCACTTCAAGTCAATTAATTTTAAACAGATATTCCTCT 2700
Qy 2701 TGTCTTAAATTTTGTGTAATATTTCTTTTGAAGTGAAGTCTCTCTGTGCCAG 2760
Db 2701 TGTCTTAAATTTTGTGTAATATTTCTTTTGAAGTGAAGTCTCTCTGTGCCAG 2760
Qy 2761 GCTGAGATACAGTGTGTGATCCAGCTCACTGACACTCCGCTCTTGAGTTTCAATGA 2820
Db 2761 GCTGAGATACAGTGTGTGATCCAGCTCACTGACACTCCGCTCTTGAGTTTCAATGA 2820
Qy 2821 TTTCTCTGCTCAAGCTTCTTAAGTGTGAGTTTACAGGACCCACCAATGCCCAGCT 2880
Db 2821 TTTCTCTGCTCAAGCTTCTTAAGTGTGAGTTTACAGGACCCACCAATGCCCAGCT 2880
Qy 2881 AATTTTGTATTTTATATAGAGAGGGGTTTCCGCAATTTGGCCAGGCTGGTCTTGAATC 2940
Db 2881 AATTTTGTATTTTATATAGAGAGGGGTTTCCGCAATTTGGCCAGGCTGGTCTTGAATC 2940
Qy 2941 CTGACGTCAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
Db 2941 CTGACGTCAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
Qy 3001 CCACTTAATTAATTTTCAATGCTATAGACATTAAGAGATTTTTCATTTTTCATGACAT 3060
Db 3001 CCACTTAATTAATTTTCAATGCTATAGACATTAAGAGATTTTTCATTTTTCATGACAT 3060
Qy 3061 TTTTCTCTCTGCAAAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
Db 3061 TTTTCTCTCTGCAAAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
Qy 3121 CATTAATATTTCTGATCATTTTCTTATCAAGAGATATATCAAGTGTGTCTCATAGA 3180
Db 3121 CATTAATATTTCTGATCATTTTCTTATCAAGAGATATATCAAGTGTGTCTCATAGA 3180
Qy 3181 ACTGCTGATTCATTTATGTTTTTCTGATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240
Db 3181 ACTGCTGATTCATTTATGTTTTTCTGATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240
Qy 3241 TCTTTGTGATTTTCACTGATTTTCAAACTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300
Db 3241 TCTTTGTGATTTTCACTGATTTTCAAACTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300
Qy 3301 GAAAAATTAATTAATAAAGACAGCCTTTTGCGCGCGAGATTTC 3345
Db 3301 GAAAAATTAATTAATAAAGACAGCCTTTTGCGCGCGAGATTTC 3345

```

```

RESULT 3
AX332252 3345 bp DNA linear PAT 09-JAN-2002
LOCUS AX332252
DEFINITION Sequence 2761 from Patent WO0194629.
ACCESSION AX332252
VERSION AX332252.1 GI:18122886
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endres, G.,
Horizgan, S., Soppet, D. R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
gene sets
Patent: WO 0194629-A 2761 13-DEC-2001;
Avalon Pharmaceuticals (US)

```

```

FEATURES
source
Location/Qualifiers
1..3345
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 951 a 746 c 720 g 928 t
ORIGIN
Query Match 100.0%; Score 3345; DB 6; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAAATCGCTCGACCACTGAATGAAAGAAAGAACTTTTACACCATTTTGTGACTTA 60
Db 1 GAAATCGCTCGACCACTGAATGAAAGAAAGAAAGAACTTTTACACCATTTTGTGACTTA 60
Qy 61 CAGAAAGAAATTTGAATTAAGAAAGAACTTATGATCTTCAAGCCATCTTCACTCTGTGT 120
Db 61 CAGAAAGAAATTTGAATTAAGAAAGAACTTATGATCTTCAAGCCATCTTCACTCTGTGT 120
Qy 121 CTTCTTATGCTTATTTGGCACTGATATGCGCAAGAGGGGAAATTTAGTGAACCCCTG 180
Db 121 CTTCTTATGCTTATTTGGCACTGATATGCGCAAGAGGGGAAATTTAGTGAACCCCTG 180
Qy 181 AAACCATGACATTTTCTATTTATGAAGGCAAGAACCGAGTCAATTTATTTCAAGTTT 240
Db 181 AAACCATGACATTTTCTATTTATGAAGGCAAGAACCGAGTCAATTTATTTCAAGTTT 240
Qy 241 AAGGCCAATCTCTCTGCTGATCTTGAATCTAATGCGGAGACAGACAAATTTTGTG 300
Db 241 AAGGCCAATCTCTCTGCTGATCTTGAATCTAATGCGGAGACAGACAAATTTTGTG 300
Qy 301 ATGAAAGGAGAGGAGCTTCTGATTTCAACAGAGCTTTGAGACAGGAAACAAATCTACT 360
Db 301 ATGAAAGGAGAGGAGCTTCTGATTTCAACAGAGCTTTGAGACAGGAAACAAATCTACT 360
Qy 361 CACAATCTCAGGTTGACAGCCCTGACGCTAATGAAATTAATGAGGGTCCAGTCCCT 420
Db 361 CACAATCTCAGGTTGACAGCCCTGACGCTAATGAAATTAATGAGGGTCCAGTCCCT 420
Qy 421 ATACCATTAAGATGAAGACATCAACAGACATGACCCAGTTTCTCCAGTCAAGATAC 480
Db 421 ATACCATTAAGATGAAGACATCAACAGACATGACCCAGTTTCTCCAGTCAAGATAC 480
Qy 481 GAAGGCTCAGTAAGGAGAACTGCGCCAGAGAAAGCCCTCTGATGCAATGCGACA 540
Db 481 GAAGGCTCAGTAAGGAGAACTGCGCCAGAGAAAGCCCTCTGATGCAATGCGACA 540
Qy 541 GACCTGATGATCCGCGCACTCCCAATGCGCAAGCTTTATACAGATTTGATCCAGCTT 600
Db 541 GACCTGATGATCCGCGCACTCCCAATGCGCAAGCTTTATACAGATTTGATCCAGCTT 600
Qy 601 CCCATATCAACATGCTATGTAATTTTCAAGTCAACAAACCGGAGCCATCTCTCTT 660
Db 601 CCCATATCAACATGCTATGTAATTTTCAAGTCAACAAACCGGAGCCATCTCTCTT 660
Qy 661 ACCCGAGAGGATCTCAGGAATTTGAATCTCTGTAAGAACTTTCCATATCTGAGATC 720
Db 661 ACCCGAGAGGATCTCAGGAATTTGAATCTCTGTAAGAACTTTCCATATCTGAGATC 720
Qy 721 TCAGTGAAGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Db 721 TCAGTGAAGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Qy 781 ATCATATGACACAGAAATTTTGAAGAGCAACAAACCTGTGAGATGAGGAGGAGGAGGAG 840
Db 781 ATCATATGACACAGAAATTTTGAAGAGCAACAAACCTGTGAGATGAGGAGGAGGAGGAG 840
Qy 841 ACTGATCTCAGCCCATCAAAATCACTGAGTGGGTGAGATGATCCGGTGAACAATAT 900
Db 841 ACTGATCTCAGCCCATCAAAATCACTGAGTGGGTGAGATGATCCGGTGAACAATAT 900
Qy 901 TCTTAGTTGACAAAGAGAGCTGCCAGATTTCCATTTTCAATTGACAGAGAGAGAT 960

```


QY	3121	CATTAAATATTCTGTACATTTTTCCTTTATCAAGAGATATATCATGTTGTCTCATAGA	3180
Db	3121	CATTAAATATTCTGTACATTTTTCCTTTATCAAGAGATATATCATGTTGTCTCATAGA	3180
QY	3181	ACTGCCCGGATTCGATTATGTTTTTTCTGATTCATCCCTGTGCCCTTATCTCTGAC	3240
Db	3181	ACTGCTGTGATTCGATTATGTTTTTTCTGATTCATCCCTGTGCCCTTATCTCTGAC	3240
QY	3241	TCCTTTGGTATTTCACTGAATTTTCAAACATTGTGTGAGAAAGAAAAAGTGAGAGCTCAG	3300
Db	3241	TCCTTTGGTATTTCACTGAATTTTCAAACATTGTGTGAGAAAGAAAAAGTGAGAGCTCAG	3300
QY	3301	GAAGAAATTAATTAATTAAGAACAGCCCTTTTGGCGCGCCGGAATTC	3345
Db	3301	GAAGAAATTAATTAATTAAGAACAGCCCTTTTGGCGCGCCGGAATTC	3345

RESULT 4				
AX410642	AX410642	3345 bp	DNA	linear
LOCUS	Sequence 3289 from Patent WO0229103.			
DEFINITION	AX410642			
ACCESSION	AX410642.1	GI:21443347		
VERSION				
				PAT 14-JUN-2002

SOURCE ORGANISM	human.	Homo sapiens	human.	Homo sapiens
1	1	1	1	1
2	2	2	2	2
3	3	3	3	3
4	4	4	4	4
5	5	5	5	5
6	6	6	6	6
7	7	7	7	7
8	8	8	8	8
9	9	9	9	9
10	10	10	10	10
11	11	11	11	11
12	12	12	12	12
13	13	13	13	13
14	14	14	14	14
15	15	15	15	15
16	16	16	16	16
17	17	17	17	17
18	18	18	18	18
19	19	19	19	19
20	20	20	20	20
21	21	21	21	21
22	22	22	22	22
23	23	23	23	23
24	24	24	24	24
25	25	25	25	25
26	26	26	26	26
27	27	27	27	27
28	28	28	28	28
29	29	29	29	29
30	30	30	30	30
31	31	31	31	31
32	32	32	32	32
33	33	33	33	33
34	34	34	34	34
35	35	35	35	35
36	36	36	36	36
37	37	37	37	37
38	38	38	38	38
39	39	39	39	39
40	40	40	40	40
41	41	41	41	41
42	42	42	42	42
43	43	43	43	43
44	44	44	44	44
45	45	45	45	45
46	46	46	46	46
47	47	47	47	47
48	48	48	48	48
49	49	49	49	49
50	50	50	50	50
51	51	51	51	51
52	52	52	52	52
53	53	53	53	53
54	54	54	54	54
55	55	55	55	55
56	56	56	56	56
57	57	57	57	57
58	58	58	58	58
59	59	59	59	59
60	60	60	60	60
61	61	61	61	61
62	62	62	62	62
63	63	63	63	63
64	64	64	64	64
65	65	65	65	65
66	66	66	66	66
67	67	67	67	67
68	68	68	68	68
69	69	69	69	69
70	70	70	70	70
71	71	71	71	71
72	72	72	72	72
73	73	73	73	73
74	74	74	74	74
75	75	75	75	75
76	76	76	76	76
77	77	77	77	77
78	78	78	78	78
79	79	79	79	79
80	80	80	80	80
81	81	81	81	81
82	82	82	82	82
83	83	83	83	83
84	84	84	84	84
85	85	85	85	85
86	86	86	86	86
87	87	87	87	

REFERENCE
1. Alvarães, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G. Gene expression profiles in liver cancer Patient: MO 0229103-A 3289 11-APR-2002; GENEX LOGIC INC (US)

source	Location/Qualifiers
1. .3345	

BASE COUNT	951 a	746 c	720 g	928 t
ORIGIN	/organism="Homo sapiens" /db_xref="taxon:9606" /note="EMBL/GenBank Accession No. U07969"			

Query Match	100.0%	Score 3345;	DB 6;	Length 3345;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 3345; Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0

Oy 1 GAATTCGCGTCCGACCACTGAATGGAAGAAAGACCTTTAAACCACCATTTTGTGACTTA 60

Db 1 GAATTCGCGTCCGACCACTGAATGGAAGAAAGACCTTTAAACCACCATTTTGTGACTTA 60

61 CAGAAAGGAAATTTGATAAAGAAACTATGATCTTCAGGGCCATCTTCACTCCCTGT 120
61 CAGAAAGGAAATTTGATAAAGAACTATGATCTTCAGGGCCATCTTCACTCCCTGT 120

122 CTTCTATGCTTATTTGGCAACTGATATGCGCCAGAGGGGAGCTTATGAGACCCCTG 180
 121 CTTCTATGCTTATTTGGCAACTGATATGCGCCAGAGGGGAGCTTATGAGACCCCTG 180

181 AAACCATGACATTTTCTATTTATGAGGCCAAGAACCGATCAATATATATTCAGTT 240

301 ATAGAACGGGAGGAGCTTCTGTATTTACACAGAGGCTTGGACAGGGAAAACAGAGTTCTATC 360
Db 241 AAGGCCAATCCTCCTGCTGAGCTTTGAATTAACGTGGGGAGACAGACAATATTTGTG 300

Db 301 ATATAACGGAGGACTTCTGTATTACAAAGAGCCTTGGACAGGAAACAAGATCTACT 360

Dy 361 CACAATCTCCAGGTTGCAGCCCTGGACGCTAATGAGATTATAGTGGAGGCTCAATCCCT 420

D**b** 361CACAACTCTCCAGGTTGCAGCCCTGACCGCTAATGGAATTATAGTGGAGGGTCCAGTCCCT 420

Oy	421	ATCACCATAGAAAGTGAAGACATTAACGACATGACCCACGTTTCTCCAGTCAAAATAC	480
Db	421	ATCACCATAGAAAGTGAAGACATTAACGACATGACCCACGTTTCTCCAGTCAAAATAC	480
Oy	481	GAAGGCTCAGTAAAGGCAGAACTCTCGCCACAGAAAGCCCTTCTTGATATGTCATGCCACA	540
Db	481	GAAGGCTCAGTAAAGGCAGAACTCTCGCCACAGAAAGCCCTTCTTGATATGTCATGCCACA	540
Oy	541	GACCTGGATGATCCGGGCACCTCCCAATAGGCGAGTTTATTACAGATATGTCATCAGCTT	600
Db	541	GACCTGGATGATCCGGGCACCTCCCAATAGGCGAGTTTATTACAGATATGTCATCAGCTT	600
Oy	601	CCCATGATCAACAATGTGATGTATCTTCAGATCAACAACAAACGGGAGGCATCTCTCTT	660
Db	601	CCCATGATCAACAATGTGATGTATCTTCAGATCAACAACAAACGGGAGGCATCTCTCTT	660

Accession	Sequence	Position
Db	661 ACCGAGAGGGATCTCAGGAATTGAATCCGCTAAGAAATCCCTCTAATATCTGGTATC	720
QY	721 TCAGTGAAGGACATGGGAGGCCAGAGTGAAGAAATTCCTTCAGTGTATACCAATCTGTGGAT	780
Db	721 TCAGTGAAGGACATGGGAGGCCAGAGTGAAGAAATTCCTTCAGTGTATATACCAATCTGTGGAT	780

781 ATCATAGTGCAGAGATATTGGAAAGCACCAAAACCTGTGGAGATGTGGAAAACTCA 840

QY 841 ACTGATCTTCACCCCATCAAAATCACTCAGGTGCGGTGATGATGCCGGTGCACAATAT 900

Db 841 ACTGATCTTCACCCCATCAAAATCACTCAGGTGCGGTGATGATGCCGGTGCACAATAT 900

Db	901	901	960
UY	TCTTAGTTCACCAAGAGAGCTG	CCCAAGATTCCCATTTTCATTG	CCAGAGAGAGAT
Db	901	901	960
UY	TCTTAGTTCACCAAGAGAGCTG	CCCAAGATTCCCATTTTCATTG	CCAGAGAGAGAT

Qy 961 ATTTACTGACTCAGCCCTTGGACCGAAGAAGATTCATATGTTTTTAAGCAGTT 1020
|||||
|||||
Db 961 ATTACTGTGACTCAGCCCTTGGACCGAAGAAGATTCATATGTTTTTAAGCAGTT 1020
|||||

Oy	1021	GCAAGAGTGA	CTACGGAAAA	CCACTTTCA	TATCCG	TGTAATTC	GTGAAAA	AGTTAAA	1080
Db	1021	GCAAGAGTGA	CTACGGAAAA	CCACTTTCA	TATCCG	TGTAATTC	GTGAAAA	AGTTAAA	1080

Db

1081 GATATTAAATGATPATACTCCACCTTACATGTCCCGTCACCAGTAACCGTATTGTGGAGTGCCAGGAG 1140

Db 1141 AATGAGACACTGGGTAA CAGTATCGGACCCCTTACTGCACATGACACGGGATGAGAAAT 1200

Db	1201	1260
1261	ACTGCCACAGTTTCTAAACATCACAGATTGTGGACCAAACTCCCAAATTCCTCCATGGAT	
1261	GGACTCTTCCATATCCAAACCTATGCTGGAATGTTACAGTAGCTATAAACAAGCTTGAAG	1320

Db	1261	GSACTCTTCCTATCCAAACCTATGCTGGAAATGTTACAGTTAAGCTAAACAGCTTGGAAG	1320
Yy	1321	AAGCAAGATACCTCCTCAGTACCAACTTAAACGATAGAGGTGTCGACAAAGATTTCAAGACC	1380

Db 1321 AAGCAGATACTCTCTCAGTACCACTTAAAGATAGAGGTGTCTGACAAAGATTTCAAGACC 1380

Qy 1381 CTTTGTTTGGCAATCAAGCTTATTGATTCATTGATCATGATCGATCCCCATCTTTGAAAA 1440

D6	1381	CTTTGTTTGTGCAAAATCAACGTTATTTGAATCAAAATGATCAGATCCCATCTTTGAAAA	1440
DY	1441	TCGATATTGAAAACCTGACTCTTGTGTAAGACACAAAACATTTGGTTCACCATCTTAAC	1500

QY 1501 ATCCAGGCCATGATGCTGATGAGCCATTCTACTGGAGTCTTAATAATTCGTATCATATC 1560

Db 1501 ATCAGGCCCTAGTGTGATGACCACTTTACTGGAGTTCTAATAATTCGTATCATATC 1560
 Qy 1561 ATAAAGGAGACAGTGAAGGACGCTTGGGGGTTGACACAGATCCCATACCAACCGGA 1620
 Db 1561 ATAAAGGAGACAGTGAAGGACGCTTGGGGGTTGACACAGATCCCATACCAACCGGA 1620
 Qy 1621 TATGTCTAATTTAAAAAGCTCTTGATTTTGAACACAGCTGTTTCCAACTTGTTC 1680
 Db 1621 TATGTCTAATTTAAAAAGCTCTTGATTTTGAACACAGCTGTTTCCAACTTGTTC 1680
 Qy 1681 AAAGCAGAAAATCCTGAGCCTCTAGTGTGTGTGAAGTACAAATGCAAGTCTTTTGGC 1740
 Db 1681 AAAGCAGAAAATCCTGAGCCTCTAGTGTGTGTGAAGTACAAATGCAAGTCTTTTGGC 1740
 Qy 1741 AAGTTACGCTTATTTGACAGATGTGAATGAGACACTGCAATTTTCCCAACAGTATTC 1800
 Db 1741 AAGTTACGCTTATTTGACAGATGTGAATGAGACACTGCAATTTTCCCAACAGTATTC 1800
 Qy 1801 CAAGCGAAAGTCAAGTGAAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1860
 Db 1801 CAAGCGAAAGTCAAGTGAAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1860
 Qy 1861 GATCCAGAAAGTCAAGTGAAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1920
 Db 1861 GATCCAGAAAGTCAAGTGAAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1920
 Qy 1921 ATTGACCAAGTCAAGTGAAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1980
 Db 1921 ATTGACCAAGTCAAGTGAAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1980
 Qy 1981 CCATATGGGGTCAAGTGAAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2040
 Db 1981 CCATATGGGGTCAAGTGAAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2040
 Qy 2041 GAGTTTCAACCTGATCTTATGATGTGAATGACAACCTCCAGAGTGAAGCAAGATAC 2100
 Db 2041 GAGTTTCAACCTGATCTTATGATGTGAATGACAACCTCCAGAGTGAAGCAAGATAC 2100
 Qy 2101 ACGGGCTTGTCTTCTGCAATCCCTCAGTGCACCTGGAAGTCTCATTTTGAAGTACT 2160
 Db 2101 ACGGGCTTGTCTTCTGCAATCCCTCAGTGCACCTGGAAGTCTCATTTTGAAGTACT 2160
 Qy 2161 GATGATGATCAGCACTTATTTGGGGGTCCTCATTTTACATTTTCCCTGCGAGTGAAGC 2220
 Db 2161 GATGATGATCAGCACTTATTTGGGGGTCCTCATTTTACATTTTCCCTGCGAGTGAAGC 2220
 Qy 2221 TTACAAAACGACTGGGAAGTTTCCAAAATCAATGGTACTCATGCCCCGACGTCTACAGG 2280
 Db 2221 TTACAAAACGACTGGGAAGTTTCCAAAATCAATGGTACTCATGCCCCGACGTCTACAGG 2280
 Qy 2281 CACAAGACTTTGAGAGAGAGGCGTATGTCTTGTATCCGATCAATGATGAGGAGTGG 2340
 Db 2281 CACAAGACTTTGAGAGAGAGGCGTATGTCTTGTATCCGATCAATGATGAGGAGTGG 2340
 Qy 2341 CCACCTTTGGAAGGATTTTCTTACAGTGTACTCTGCAATTTTCCCTGCGAGTGAAGC 2400
 Db 2341 CCACCTTTGGAAGGATTTTCTTACAGTGTACTCTGCAATTTTCCCTGCGAGTGAAGC 2400
 Qy 2401 TGTTCGCGGAGAGGATCACAAGCTGGGATACCACTGCGGATGGAGTGGTATTA 2460
 Db 2401 TGTTCGCGGAGAGGATCACAAGCTGGGATACCACTGCGGATGGAGTGGTATTA 2460
 Qy 2461 CTGCTGACCACTTCTGAGTATGTTGATTAATTTTGAAGTGTGTATTCGCAATAAG 2520
 Db 2461 CTGCTGACCACTTCTGAGTATGTTGATTAATTTTGAAGTGTGTATTCGCAATAAG 2520
 Qy 2521 AAGGATTAAGGCAAGATTAATTTGAAAGTGTCTCAAGATCTGAAGTAAACCTCTGGA 2580
 Db 2521 AAGGATTAAGGCAAGATTAATTTGAAAGTGTCTCAAGATCTGAAGTAAACCTCTGGA 2580
 Qy 2581 AGCTGAATTTGAAAAGATTTGAAATTTATATGCAAGTGTCTTACAGCAACAACA 2640
 Db 2581 AGCTGAATTTGAAAAGATTTGAAATTTATATGCAAGTGTCTTACAGCAACAACA 2640

Db 2581 AGCTGAATTTGAAAAGATTTGAAATTTATATGCAAGTGTCTTACAGCAACAACA 2640
 Qy 2641 TCTCATCCATTAATCTTATGATCAATGATGATTAATTTTAAACAGATATTCCTCT 2700
 Db 2641 TCTCATCCATTAATCTTATGATCAATGATGATTAATTTTAAACAGATATTCCTCT 2700
 Qy 2701 TGTCTTAAATATTTGCTAAATATTTCTTTTGAAGTGAAGTCTGTCTGTGCGCCAG 2760
 Db 2701 TGTCTTAAATATTTGCTAAATATTTCTTTTGAAGTGAAGTCTGTCTGTGCGCCAG 2760
 Qy 2761 GCTGAGTCAAGTGTGTGATCCAGCTCACTGCAACCTCCGCTCTGGGTTCAATGA 2820
 Db 2761 GCTGAGTCAAGTGTGTGATCCAGCTCACTGCAACCTCCGCTCTGGGTTCAATGA 2820
 Qy 2821 TTTCTGCTGCTGAGCTTCCATGATGCTGGGTTTACAGGACCAACCATCCAGCT 2880
 Db 2821 TTTCTGCTGCTGAGCTTCCATGATGCTGGGTTTACAGGACCAACCATCCAGCT 2880
 Qy 2881 AATTTTGTATTTTAAATAGACAGGGGTTTCCGCAATTTGGCCAGGCTGTGGAATC 2940
 Db 2881 AATTTTGTATTTTAAATAGACAGGGGTTTCCGCAATTTGGCCAGGCTGTGGAATC 2940
 Qy 2941 CTGAGCTCAAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
 Db 2941 CTGAGCTCAAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
 Qy 3001 CCTACTGATATTTGATGATGCTGATGATGATGATGATGATGATGATGATGATGAT 3060
 Db 3001 CCTACTGATATTTGATGATGCTGATGATGATGATGATGATGATGATGATGATGAT 3060
 Qy 3061 TTTTCTCTCTGCAAAATGCTTACGATGATGATGATGATGATGATGATGATGATGAT 3120
 Db 3061 TTTTCTCTCTGCAAAATGCTTACGATGATGATGATGATGATGATGATGATGATGAT 3120
 Qy 3121 CATTAATATTTCTGATCAATTTTCTTTATCAAGAGATATATCATGTTGTCTATGA 3180
 Db 3121 CATTAATATTTCTGATCAATTTTCTTTATCAAGAGATATATCATGTTGTCTATGA 3180
 Qy 3181 ACTGCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
 Db 3181 ACTGCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
 Qy 3241 TCCCTTGTATTTTCACTGATTTTCAAAATTTTCTGAGAGAGAGAGAGAGAGAGAG 3300
 Db 3241 TCCCTTGTATTTTCACTGATTTTCAAAATTTTCTGAGAGAGAGAGAGAGAGAGAG 3300
 Qy 3301 GAAAAATTAATTAATTAAGAGACAGCTTTTGGCGCGCAATTC 3345
 Db 3301 GAAAAATTAATTAATTAAGAGACAGCTTTTGGCGCGCAATTC 3345
 RESULT 5
 HSDU7969
 LOCUS
 DEFINITION
 Human intestinal peptide-associated transporter HPT-1 mRNA,
 complete cds.
 U07969.1 GI:483391
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 3345)
 Dantzig, A.H., Hoskins, J., Tabas, L.B., Bright, S., Shepard, R.L.,
 Jenkins, I.L., Duckworth, D.C., Sportsman, R., Mackensen, D.,
 Roelcke, P.R. and Skatrud, P.L.
 Association of intestinal peptide transport with a protein related
 to the cadherin superfamily
 JOURNAL
 Science 264 (5157), 430-433 (1994)
 MEDLINE
 94204643
 PUBMED
 8153632
 REFERENCE
 2 (bases 1 to 3345)

AUTHORS
Rosteck, P. R.
Direct Submission
Submitted (24-MAR-1994) Paul R. Rosteck Jr., Lilly Research
JOURNALS Laboratories, Eli Lilly and Co., Lilly Corporate Center,
Indianapolis, IN 46285, USA

FEATURES
Location/Qualifiers

SOURCE
1. .3345
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Caco-2"
/cell_type="adenocarcinoma"
/tissue_type="colon"
88. .2586
/product="intestinal peptide-associated transporter HPT-1"
/protein_id="AA19021.1"
/db_xref="GI:483392"
/translation="MILQAHLSICLIMLYATGYGQSGKPSGLKPMPSIYEGEP
SOIIPQKAMPVATFELTGETDNIPIREGLLYNRALDRETRSHNIOVALDAN
GIIVEGVPTIEVKDINDRPTLOSKEYSVONRPGKPLRYNATLDLPATN
GOLOYOIVOLPMINMYFOINNKTKGALSITRESGQELNPARNSYNIIVSYDMG
OSNSPSTDSVDIIVTENTMKA PKVEMENSTDPPICTOVRMNDPEAGSINDK
EKIPRPESIDOBSDIYTOPLDREKDAYFYAVADERYGKPLSPLETHVAKDIN
DNPPICSPVTFVEVOENRNLNSIGLTAHDDEENTANSFLNRYLVEQTPLEMDG
LFLIQYAGMLQAKOSLKKQDTPQVNLTFEVDKDFKTLCPQVINVIDINDOIPIFE
KSDYGNLTLEADNIGSTLIOTADEDEPTSSKILYHI IKDSEGRGVTDPH
NTGVILIKKPLDEPETAIVSNIVFAENRPELVGVKNASSPAKFTLLIYDVNAPQF
SOHFOAKYSEVDVATGKGNMVAKDEGLDISYSLRDPKMLKIDHTVGEI.FSVAP
LDREAGSPRYOVAVATVGGSSISVSERFLIIMDVNDNPKRLAKDTG.FPCHPLSA
PGLIFPATDDQHLERPHPTFSLSGSIQNDMEVSKINGTARSTRTDDEERY
VLIIRINDGRPLBGLVSLPVTSCSEVSCFRPAHQGTITVGVAVGILLTTLIV
IGIILAVFRIKDKGKDNESAQSEVPSPLRS"

CDS
BASE COUNT 951 a 746 c 720 g 928 t

Query Match 100.0%; Score 3345; DB 9; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; **Mismatches** 0; **Indels** 0; **Gaps** 0;

QY	1	GAATTCGCTCCACCACTGAATGGAAGAAAGACTTTTAAACCACTTTTGTGACTTA	60
DB	1	GAATTCGCTCCACCACTGAATGGAAGAAAGACTTTTAAACCACTTTTGTGACTTA	60
QY	61	CAGAAAGAAATTTGAATTAAGAAAACCTATGATCTTCAGGCCCATCTTCACTCCGTGT	120
DB	61	CAGAAAGAAATTTGAATTAAGAAAACCTATGATCTTCAGGCCCATCTTCACTCCGTGT	120
QY	121	CTTCTTATGCTTTATTTGGCACTGATATGCCCAAGAGGGGAAATTAGTGAACCCCTG	180
DB	121	CTTCTTATGCTTTATTTGGCACTGATATGCCCAAGAGGGGAAATTAGTGAACCCCTG	180
QY	181	AAACCCATGACATTTTCTATTTATGAAAGGCCAAGAACCGAGTCAATTAATTCAGATT	240
DB	181	AAACCCATGACATTTTCTATTTATGAAAGGCCAAGAACCGAGTCAATTAATTCAGATT	240
QY	241	AAGGCCAATCTCTGCTGTGATTTGAACCTAAGTGGGAGACAGACAACTATTTGTTG	300
DB	241	AAGGCCAATCTCTGCTGTGATTTGAACCTAAGTGGGAGACAGACAACTATTTGTTG	300
QY	301	ATAGAACGGAGGAGACTTCTGTATTAACAAGACCTTTGACAGGAAACAAATCTACT	360
DB	301	ATAGAACGGAGGAGACTTCTGTATTAACAAGACCTTTGACAGGAAACAAATCTACT	360
QY	361	CAGAACTCCAGGTTGAGCCCTGAGCGCTAATGGAATTAAGTGAAGGCTCCAGTCCCT	420
DB	361	CAGAACTCCAGGTTGAGCCCTGAGCGCTAATGGAATTAAGTGAAGGCTCCAGTCCCT	420
QY	421	ATCACATAGAAAGTGAAGACATCAAGCAATCGAACCCAGTTTCTCCAGTCAAGATAC	480
DB	421	ATCACATAGAAAGTGAAGACATCAAGCAATCGAACCCAGTTTCTCCAGTCAAGATAC	480
QY	481	GAAGGCTCAGTAAAGGAGAACTCTCGCCAGGAAAGCCCTTCTTGTATGTCAATGCCACA	540
DB	481	GAAGGCTCAGTAAAGGAGAACTCTCGCCAGGAAAGCCCTTCTTGTATGTCAATGCCACA	540

DB	481	GAAGGCTCAGTAAAGGAGAACTCTCGCCAGGAAAGCCCTTCTTGTATGTCAATGCCACA	540
QY	541	GACCTGGATATCCGGCACTCCCAATGGCCAGCTTTATTAACAAGATTCACCGAGCT	600
DB	541	GACCTGGATATCCGGCACTCCCAATGGCCAGCTTTATTAACAAGATTCACCGAGCT	600
QY	601	CCCATGATCAACATGTCATGATCTTTCAGATCAACAACAAAGGAGCCATCTCTCT	660
DB	601	CCCATGATCAACATGTCATGATCTTTCAGATCAACAACAAAGGAGCCATCTCTCT	660
QY	661	ACCCGAGAGGAGATCTCAGAAATTTGAATTCCTGCTAAGAAATCTTCTAATATCTGTGATC	720
DB	661	ACCCGAGAGGAGATCTCAGAAATTTGAATTCCTGCTAAGAAATCTTCTAATATCTGTGATC	720
QY	721	TCAGTAAAGCAATGGAGGCCAGATGGAATTCCTTCAGTATACCAATCTGTGAT	780
DB	721	TCAGTAAAGCAATGGAGGCCAGATGGAATTCCTTCAGTATACCAATCTGTGAT	780
QY	781	ATCATAGTGCACAGAAATATTTGAAAGCACAAACCTGTGAGATGTGGAACAACTCA	840
DB	781	ATCATAGTGCACAGAAATATTTGAAAGCACAAACCTGTGAGATGTGGAACAACTCA	840
QY	841	ACTGATCTTCAACCCCATCAAAATCACTCAGGTGCGGTGAATGATCCCGGTGACAAATAT	900
DB	841	ACTGATCTTCAACCCCATCAAAATCACTCAGGTGCGGTGAATGATCCCGGTGACAAATAT	900
QY	901	TCCTTGTGTCACAAAGAAAGCTGCGCAAGATTTCCCATTTTCAATTGACCAAGAAAGAT	960
DB	901	TCCTTGTGTCACAAAGAAAGCTGCGCAAGATTTCCCATTTTCAATTGACCAAGAAAGAT	960
QY	961	ATTTAAGTGCATGAGCCCTTGGACCGAGAAAGAAAGATGATGTTTATGACGTT	1020
DB	961	ATTTAAGTGCATGAGCCCTTGGACCGAGAAAGAAAGATGATGTTTATGACGTT	1020
QY	1021	GCAAGAGATGAGTACGAGAAACCACTTTCATATCCGCTGAAATTCATGTAAGTTAA	1080
DB	1021	GCAAGAGATGAGTACGAGAAACCACTTTCATATCCGCTGAAATTCATGTAAGTTAA	1080
QY	1081	GATATTAATGATATATCACCTCATGTCCGTACCAAGTAAACCGTATTTGAGGTCCAGGAG	1140
DB	1081	GATATTAATGATATATCACCTCATGTCCGTACCAAGTAAACCGTATTTGAGGTCCAGGAG	1140
QY	1141	AATGAACGACTGGGTAAAGATTCGAGACCTTACGACATGACAGGAGTGAAGAAAT	1200
DB	1141	AATGAACGACTGGGTAAAGATTCGAGACCTTACGACATGACAGGAGTGAAGAAAT	1200
QY	1201	ACTGCCAACAGTTTCTTAAACTAAGAGATTGGAGCAAAATCCCAAATCCCATGAT	1260
DB	1201	ACTGCCAACAGTTTCTTAAACTAAGAGATTGGAGCAAAATCCCAAATCCCATGAT	1260
QY	1261	GGAATCTTCTTATCCAAACCTATGCTGGAATGTTACAGTTCATGCTTGAAG	1320
DB	1261	GGAATCTTCTTATCCAAACCTATGCTGGAATGTTACAGTTCATGCTTGAAG	1320
QY	1321	AAGCAAGATATCTTCAGTACAACTTAAGATAGAGGTCTGACAAAGATTTCAAGACC	1380
DB	1321	AAGCAAGATATCTTCAGTACAACTTAAGATAGAGGTCTGACAAAGATTTCAAGACC	1380
QY	1381	CTTGTGTTTGTGCAATCAACGTTATTTGATATCAATGATCAAGTCCCATCTTTGAAAA	1440
DB	1381	CTTGTGTTTGTGCAATCAACGTTATTTGATATCAATGATCAAGTCCCATCTTTGAAAA	1440
QY	1441	TCAGATTATGAGAACTGACTCTTGTGTAAGACAAACATTTGGGTCCACATCTTAACC	1500
DB	1441	TCAGATTATGAGAACTGACTCTTGTGTAAGACAAACATTTGGGTCCACATCTTAACC	1500
QY	1501	ATCCAGGACATGATGCTGATGAGCCATTTATCGGAGTTCTTAAATCTGTATCATATC	1560
DB	1501	ATCCAGGACATGATGCTGATGAGCCATTTATCGGAGTTCTTAAATCTGTATCATATC	1560
QY	1561	ATTAAGGAGACAGTGAAGGAGCGCTGGGGTTGACACAGATCCCATACCAACCGGA	1620
DB	1561	ATTAAGGAGACAGTGAAGGAGCGCTGGGGTTGACACAGATCCCATACCAACCGGA	1620

OY		2701	TGTCCTTAATATTTGCTAAATAAATTTCTTTTTTGGAGGTGAAGCTTGCTGCACGACCAG	2760
Db		2701	TGTCCTTAATATTTGCTAAATAATTTCTTTTTTGGAGGTGAAGCTTGCTGCACGACCAG	2760
OY		2761	GCTGAGATCAAGTGTGTGATGCCAGCTCATCTGCACACTCCGCTCCTCGGATTCACATGA	2820
Db		2761	GCTGAGATCAAGTGTGTGATGCCAGCTCATCTGCACACTCCGCTCCTCGGATTCACATGA	2820
OY		2821	TTCTCTGCTCCAGCTTCTCPAAGTAGCTGGGTTTACAGGCACCAACAATGCGCCAGCT	2880
Db		2821	TTCTCTGCTCCAGCTTCTCPAAGTAGCTGGGTTTACAGGCACCAACAATGCGCCAGCT	2880
OY		2881	AATTTTGGATTTTAAADAGAAGAGGGGTTTCGCCATTTGGCCAGGCTGGCTTGAATCTC	2940
Db		2881	AATTTTGGATTTTAAADAGAAGAGGGGTTTCGCCATTTGGCCAGGCTGGCTTGAATCTC	2940
OY		2941	CTGACGTCAGATGATCTGCCTGCTGGTCTCCCAATACAGGACATGAACCATGCAACCA	3000
Db		2941	CTGACGTCAGATGATCTGCCTGCTGGTCTCCCAATACAGGACATGAACCATGCAACCA	3000
OY		3001	CCATCTTAGATATTTCATGTGCTATPACACATTAGAGAGATTTTTCATTTTCCATGACAT	3060
Db		3001	CCATCTTAGATATTTCATGTGCTATPACACATTAGAGAGATTTTTCATTTTCCATGACAT	3060
OY		3061	TTTTCTCTCTCGCAAAATGGCTTACTACTGCTGTGTTTTCCCTTTGGGGGCAAGCAGACT	3120
Db		3061	TTTTCTCTCTCGCAAAATGGCTTACTACTGCTGTGTTTTCCCTTTGGGGGCAAGCAGACT	3120
OY		3121	CATTAAATATTCTGTACATTTTTCTTTATCAAGAGATATATCATGTTGTCTCATAGA	3180
Db		3121	CATTAAATATTCTGTACATTTTTCTTTATCAAGAGATATATCATGTTGTCTCATAGA	3180
OY		3181	ACTGCTGATTCATTTGATTTTCTGANTTCATCTCTGTGTCCTTTCATCTCTTGAC	3240
Db		3181	ACTGCTGATTCATTTGATTTTCTGANTTCATCTCTGTGTCCTTTCATCTCTTGAC	3240
OY		3241	TCCTTTGGTATTTTCACTGAAATTTTCAAACTTTGTCAAGAGAAAAGTAGGAGACTCAG	3300
Db		3241	TCCTTTGGTATTTTCACTGAAATTTTCAAACTTTGTCAAGAGAAAAGTAGGAGACTCAG	3300
OY		3301	GAAAAATTAATTAATTAATAAAGAACAGCCTTTTGCAGCGCGCGAATTC	3345
Db		3301	GAAAAATTAATTAATTAATAAAGAACAGCCTTTTGCAGCGCGCGAATTC	3345
RESULT 6				
AX330312				
LOCUS	AX330312	3654 bp	DNA	linear PAT 09-JAN-2002
DEFINITION	Sequence 821 from Patent WO0194629.			
ACCESSION	AX330312			
VERSION	AX330312.1	GI:18103290		
KEYWORDS				
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE				
AUTHORS	1 Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,			
	Horrigan,S., Soppet,D.R. and Weaver,Z.			
TITLE	Cancer gene determination and therapeutic screening using signature			
JOURNAL	gene sets			
	Patent: WO 0194629-A 821.13-DEC-2001;			
FEATURES	Avalon Pharmaceuticals (US)			
source	location/Qualifiers			
	1..3654			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
BASE COUNT	1052 a 804 c 777 g 1021 t			
ORIGIN				
Query Match	99.1%; Score 3314.8; DB 6; Length 3654;			
Best Local Similarity	99.8%; Pred. No. 0;			
Matches 3319; Conservative	0; Mismatches 7; Indels 0; Gaps 0;			

QY 8 GTCTGACCACTGAATGGAAGAAAAGACTTTTAACCAACATTTTGTGACTTACAGAAAG 67
Db 13 GTCTGACCACTGAATGGAAGAAAAGACTTTTAACCAACATTTTGTGACTTACAGAAAG 72
QY 68 GAATTTGAATTAAGAAAAGCTATGATTACTTACAGGCCCATCTTCACTCCCTGTGTCTTTA 127
Db 73 GAATTTGAATTAAGAAAAGCTATGATTACTTACAGGCCCATCTTCACTCCCTGTGTCTTTA 132
QY 128 TGCTTTATTTGGCAACTGATATGSCCAAGAGGGGAAAGTTTAGTGAAGCCCTGAAACCA 187
Db 133 TGCTTTATTTGGCAACTGATATGSCCAAGAGGGGAAAGTTTAGTGAAGCCCTGAAACCA 192
QY 188 TGAATTTTCTATTTATTAAGAGCCCAAGACCGAGTCAATTAATTAATTCAGTTTAAGSCCA 247
Db 193 TGACATTTTCTATTTATTAAGAGCCCAAGACCGAGTCAATTAATTAATTCAGTTTAAGSCCA 252
QY 248 ATCTCTCTGTGTGACTTTTGAATTAATCTGGGGAGACAGACACATATTTGTGATGAAC 307
Db 253 ATCTCTCTGTGTGACTTTTGAATTAATCTGGGGAGACAGACACATATTTGTGATGAAC 312
QY 308 GGGAGGGAGCTTCTGTATTAACAAGAGCTTTGACAGGGAAACAAGATCTATCAACATC 367
Db 313 GGGAGGGAGCTTCTGTATTAACAAGAGCTTTGACAGGGAAACAAGATCTATCAACATC 372
QY 368 TCCAGGTTGAGCCCTGAGAGCTTAATGAAATTAATGAGAGGTCAGTCCCTATCAACA 427
Db 373 TCCAGGTTGAGCCCTGAGAGCTTAATGAAATTAATGAGAGGTCAGTCCCTATCAACA 432
QY 428 TAGAATGGAAGACATCAAGACATGACACCGTTCCTCACTCAAGTACAGAGGCT 487
Db 433 TAAAGTGAAGGACATCAAGACATGACACCGTTCCTCACTCAAGTACAGAGGCT 492
QY 488 CAGTAAGGAGCACTCTCGCCAGGAAAGCCCTTCTGTATGTCAAAGCCACAGACCTGG 547
Db 493 CAGTAAGGAGCACTCTCGCCAGGAAAGCCCTTCTGTATGTCAAAGCCACAGACCTGG 552
QY 548 ATGATCCGGCACTCCCAATGAGCACTTTATTAACAGATTTGTCACAGTCCAGTCCCATGA 607
Db 553 ATGATCCGGCACTCCCAATGAGCACTTTATTAACAGATTTGTCACAGTCCAGTCCCATGA 612
QY 608 TCAACAAATGTCATGATCTTTCAAGTCAACAAACAAACGAGAGCCATCTTACCCGAG 667
Db 613 TCAACAAATGTCATGATCTTTCAAGTCAACAAACAAACGAGAGCCATCTTACCCGAG 672
QY 668 AGGATATCCAGAAATTTGAATCTCTGCTAAGAAATCTTCTTAATCTGTGATCTCAAGTGA 727
Db 673 AGGATATCCAGAAATTTGAATCTCTGCTAAGAAATCTTCTTAATCTGTGATCTCAAGTGA 732
QY 728 AGGACATGGAGGAGCAGAGTGAAGAAATCTTCAAGTATCAACATCTGTGATATCATAG 787
Db 733 AGGACATGGAGGAGCAGAGTGAAGAAATCTTCAAGTATCAACATCTGTGATATCATAG 792
QY 788 TGAACAGAAATTTTGGAAAGACCAAAACCTGTGAGATGTGAAAACTCAACCTGATC 847
Db 793 TGAACAGAAATTTTGGAAAGACCAAAACCTGTGAGATGTGAAAACTCAACCTGATC 852
QY 848 CTACCCCATCAAAATCACTCAGGTGAGTGAATGATCCCGGTGCACAATATTCCTTAG 907
Db 853 CTACCCCATCAAAATCACTCAGGTGAGTGAATGATCCCGGTGCACAATATTCCTTAG 912
QY 908 TTGACAAAGAGAGCTGCAAGATTTCCATTTTCAATTGACAGAGAGAGATATTTAGC 967
Db 913 TTGACAAAGAGAGCTGCAAGATTTCCATTTTCAATTGACAGAGAGAGATATTTAGC 972
QY 968 TGAATCAAGCCCTTGGACCGAGAAAGAAAGATGATATGTTTTTTATGACAGTTGCAAGG 1027
Db 973 TGAATCAAGCCCTTGGACCGAGAAAGAAAGATGATATGTTTTTTATGACAGTTGCAAGG 1032
QY 1028 ATGAGTACGGAAGAAACCACTTTCAATCCGCTGGAATTTCAATGTAAGTTAAAGTAAATTA 1087
Db 1033 ATGAGTACGGAAGAAACCACTTTCAATCCGCTGGAATTTCAATGTAAGTTAAAGTAAATTA 1092

QY 1088 ATGATTAATCCACCTACATGTCGGTCAACAGTAAACCGTATTTAGGTCCAGAGATGTAAC 1147
Db 1093 ATGATTAATCCACCTACATGTCGGTCAACAGTAAACCGTATTTAGGTCCAGAGATGTAAC 1152
QY 1148 GACTGGGTAAACAGTATGGGACCTTTATCTGACATGACAGAGATGAAGAAATATCTGCA 1207
Db 1153 GACTGGGTAAACAGTATGGGACCTTTATCTGACATGACAGAGATGAAGAAATATCTGCA 1212
QY 1208 ACAATTTTCTAAACCTACAGATTTGTGGAGCAATCCCAAACTTCCATGGATGGAATCT 1267
Db 1213 ACAATTTTCTAAACCTACAGATTTGTGGAGCAATCCCAAACTTCCATGGATGGAATCT 1272
QY 1268 TCTTAATCCAAACCTATGCTGGAATGTTACAGTTAGCTTAAACAGTCCCTGAAGAGCAAG 1327
Db 1273 TCTTAATCCAAACCTATGCTGGAATGTTACAGTTAGCTTAAACAGTCCCTGAAGAGCAAG 1332
QY 1328 ATACTCTCTCAATCACTTAAGATAGAGGTGTCTGCAAAAGATTTCAAGACCTTTGTT 1387
Db 1333 ATACTCTCTCAATCACTTAAGATAGAGGTGTCTGCAAAAGATTTCAAGACCTTTGTT 1392
QY 1388 TTGTGAAATCAACGTTATTTGATATCAATGATCCCATCTTTGAAAAATCAAGAT 1447
Db 1393 TTGTGAAATCAACGTTATTTGATATCAATGATCCCATCTTTGAAAAATCAAGAT 1452
QY 1448 ATGAAAACCTGACTCTTGTGTAAGACACAACAATGGGTCCACCATCTTAACCATCAGG 1507
Db 1453 ATGAAAACCTGACTCTTGTGTAAGACACAACAATGGGTCCACCATCTTAACCATCAGG 1512
QY 1508 CCACTGATGCTGATAGACCATTTACTGGGAGTTCTAAATTTCTGTATCATATCAATAAAG 1567
Db 1513 CCACTGATGCTGATAGACCATTTACTGGGAGTTCTAAATTTCTGTATCATATCAATAAAG 1572
QY 1568 GAGACAGTGAAGGAGCCCTGGGGGTTGACACATCCCATACCAACCCGATATGTCA 1627
Db 1573 GAGACAGTGAAGGAGCCCTGGGGGTTGACACATCCCATACCAACCCGATATGTCA 1632
QY 1628 TAAATTAAGAAAGCCTCTTGTGATTTTGAAGACAGACCTGTTTCCAAATTTGTGAAAGCAG 1687
Db 1633 TAAATTAAGAAAGCCTCTTGTGATTTTGAAGACAGACCTGTTTCCAAATTTGTGAAAGCAG 1692
QY 1688 AAAATCTGAGCCTCTAGTGTGTTGTGTGAAAGTACAAATGCAAGTCTTTTGGCAAGTTCA 1747
Db 1693 AAAATCTGAGCCTCTAGTGTGTTGTGTGAAAGTACAAATGCAAGTCTTTTGGCAAGTTCA 1752
QY 1748 CGCTTAATGTAAGAAATGTAAGTGAAGACACCTGAATTTTCCCAACGATTTCCAGGCA 1807
Db 1753 CGCTTAATGTAAGAAATGTAAGTGAAGACACCTGAATTTTCCCAACGATTTCCAGGCA 1812
QY 1808 AAGTCAGTGAAGATGATATATGAGCACTAAAGTGGGCAATGATGCTGCCAAGATTCAG 1867
Db 1813 AAGTCAGTGAAGATGATATATGAGCACTAAAGTGGGCAATGATGCTGCCAAGATTCAG 1872
QY 1868 AAGTCTGACATTAACCTATTCATGAGGGAGACACAAGAGGTTGGCTTAAATTTGACC 1927
Db 1873 AAGTCTGACATTAACCTATTCATGAGGGAGACACAAGAGGTTGGCTTAAATTTGACC 1932
QY 1928 ACGTGAATGCTGATGATCTTAAGTGTGCTCCATTTGACAGAGAAAGCCGAAAGTCCATATC 1987
Db 1933 ACGTGAATGCTGATGATCTTAAGTGTGCTCCATTTGACAGAGAAAGCCGAAAGTCCATATC 1992
QY 1988 GGGTACAGATGATGAGCACAAGAAATGAGGGGGTCTTCTTAAAGCTCTGTGTCAAGATTC 2047
Db 1993 GGGTACAGATGATGAGCACAAGAAATGAGGGGGTCTTCTTAAAGCTCTGTGTCAAGATTC 2052
QY 2048 ACCTGATCTTATGATGATGATGATCAACCTTCCAGGCTAGCCAGACATACACGGGCT 2107
Db 2053 ACCTGATCTTATGATGATGATGATCAACCTTCCAGGCTAGCCAGACATACACGGGCT 2112
QY 2108 TGTTCCTTGGCAATCCCTCAAGTGCACCTGGAAGTCTCATTTTCCAGGCTACATGATGATG 2167
Db 2113 TGTTCCTTGGCAATCCCTCAAGTGCACCTGGAAGTCTCATTTTCCAGGCTACATGATGATG 2172
QY 2168 ATCAGCACTTAATTCGGGGTCCCATTTTACATTTTCCCTCGGCAAGTGAAGCTTACAA 2227

Db 2173 ATGAGCATTAATTTGGGGTCCCATTTTACATTTTCCCTGGGAGTGGAAAGCTTACAA 2232
 Qy 2228 ACGACTGGGAAGTTTCCMAATCAATGTACTACGCCGACCTGTCTACAGGACACAG 2287
 Db 2233 ACGACTGGGAAGTTTCCMAATCAATGTACTACGCCGACCTGTCTACAGGACACAG 2292
 Qy 2288 ACTTTGAGGAGAGGGCGTATGTCTGTGATCCGCATCATGATGGGGGTCCGCCACCT 2347
 Db 2293 AGTTTGGAGAGAGGAGTATGTCTGTGATCCGCATCATGATGGGGGTCCGCCACCT 2352
 Qy 2348 TGGAGGCAATTTGTTCTTTACCACTTACATCTCAGTTGTGTGAAGAAAGTTGTTCC 2407
 Db 2353 TGGAGGCAATTTGTTCTTTACCACTTACATCTCAGTTGTGTGAAGAAAGTTGTTCC 2412
 Qy 2408 GGGCAGCAGAGTCCACAGACTGGGATACCACCTGTGGGACATGGCAGTTGTAATCTGCTGA 2467
 Db 2413 GGGCAGCAGAGTCCACAGACTGGGATACCACCTGTGGGACATGGCAGTTGTAATCTGCTGA 2472
 Qy 2468 CCAACCTTCTGGTATGTAATGTAATTTTATAGAGTTGTGTTATCCGATAAAGAGATA 2527
 Db 2473 CCACCTTCTGGTATGTAATGTAATTTTATAGAGTTGTGTTATCCGATAAAGAGATA 2532
 Qy 2528 AAGGCAAGATTAATGTTGAAGTCTCAGCATCTGAAGTCAAACTCTGAGAAAGCTGAA 2587
 Db 2533 AAGGCAAGATTAATGTTGAAGTCTCAGCATCTGAAGTCAAACTCTGAGAAAGCTGAA 2592
 Qy 2588 TTTGAAAAGAAATTTGAATTTATATAGCAAGTCTATTTCAGCAACAACATCTCATC 2647
 Db 2593 TTTGAAAAGAAATTTGAATTTATATAGCAAGTCTATTTCAGCAACAACATCTCATC 2652
 Qy 2648 CTATTACTTTTCACTAACGCGCATTAATTTTAAACGAAATTCCTCTGTGCTT 2707
 Db 2653 CTATTACTTTTCACTAACGCGCATTAATTTTAAACGAAATTCCTCTGTGCTT 2712
 Qy 2708 TAATATTGCTAAATATTCTTTTGTGAGTGAAGTCTGTGCTGTGGCCAGGCTGGAG 2767
 Db 2713 TAATATTGCTAAATATTCTTTTGTGAGTGAAGTCTGTGCTGTGGCCAGGCTGGAG 2772
 Qy 2768 TACAGTGTGTGATCCAGCTCACTGCAACCTCGGCTCTGTGGTTCAATGATTTCTCT 2827
 Db 2773 TACAGTGTGTGATCCAGCTCACTGCAACCTCGGCTCTGTGGTTCAATGATTTCTCT 2832
 Qy 2828 GCGTCACTTCTTAAGTGAAGTGGGTTTACAGGACCAACCACTGCGCAGCTAATTTT 2887
 Db 2833 GCGTCACTTCTTAAGTGAAGTGGGTTTACAGGACCAACCACTGCGCAGCTAATTTT 2892
 Qy 2888 GTATTTTAAATAGAGACGGGGTTTCGCCATTTGGCCAGGCTGTGTAATCTCTGACGT 2947
 Db 2893 GTATTTTAAATAGAGACGGGGTTTCGCCATTTGGCCAGGCTGTGTAATCTCTGACGT 2952
 Qy 2948 CAAATGATCTGCTGCTGTGTCTCCCAATACAGGACATGAACAACCTGACCCACTTACT 3007
 Db 2953 CAAATGATCTGCTGCTGTGTCTCCCAATACAGGACATGAACAACCTGACCCACTTACT 3012
 Qy 3008 AGAATTTCAATGCTATAGACATTTAGAGATTTTCAATTTTCCATGACATTTTCTCT 3067
 Db 3013 AGAATTTCAATGCTATAGACATTTAGAGATTTTCAATTTTCCATGACATTTTCTCT 3072
 Qy 3068 CTGCGCAAAATGGCTAGCTACTGTGTTTCCCTTTGGGGCAAGACAGACTAATTA 3127
 Db 3073 CTGCGCAAAATGGCTAGCTACTGTGTTTCCCTTTGGGGCAAGACAGACTAATTA 3132
 Qy 3128 TAATCTGATCAATTTTCTTTATCAAGAGATATATCATGTGTCTCATAGAACTGCTT 3187
 Db 3133 TAATCTGATCAATTTTCTTTATCAAGAGATATATCATGTGTCTCATAGAACTGCTT 3192
 Qy 3188 GGATTCATTTATGTTTCTGATTCATCTGTGTCTCTTCACTCTTCACTCTTGG 3247
 Db 3193 GGATTCATTTATGTTTCTGATTCATCTGTGTCTCTTCACTCTTCACTCTTGG 3252
 Qy 3248 GTATTTCACTGAATTTCAACATTTGTGAGAGAGAAAAAGTGAAGCTCAGGAAAAAT 3307

Db 3253 GTATTTCACTGAATTTCAACATTTGTGACAGAGAGAAAAACGTGAGACTCAGAAAAAT 3312
 Qy 3308 AATTAATTAAGAAACAGGCTTTTGC 3333
 Db 3313 AATTAATTAAGAAACAGGCTTTTTC 3338
 RESULT 7
 AX335495
 LOCUS
 DEFINITION
 Sequence 6004 from Patent WO0194629.
 AX335495
 VERSION
 AX335495.1 GI:18126214
 KEYWORDS
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1
 AUTHORS
 Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endres, G.,
 Horrigan, S., Soppet, D. R. and Weaver, Z.
 TITLE
 Cancer gene determination and therapeutic screening using signature
 gene sets
 JOURNAL
 Patent: WO 0194629-A 6004 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES
 source
 1..3654
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT
 1052 a 804 c 777 g 1021 t
 ORIGIN
 Query Match 99.1%; Score 3314.8; DB 6; Length 3654;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 3319; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Qy 8 GTCTGACCACTGAATGGAAGAAAGACTTTTACCACATTTTGTACTTACAGAAAG 67
 Db 13 GTCTGACCACTGAATGGAAGAAAGACTTTTACCACATTTTGTACTTACAGAAAG 72
 Qy 68 GAATTTGAATTAAGAAACTATGATATCTTACGCCCATCTTCACTCCTGTGCTTTA 127
 Db 73 GAATTTGAATTAAGAAACTATGATATCTTACGCCCATCTTCACTCCTGTGCTTTA 132
 Qy 128 TGCTTTATTTGGCAATCGATATAGGCCAAGAGGGAAGTTTATGAGACCCCTGAACCA 187
 Db 133 TGCTTTATTTGGCAATCGATATAGGCCAAGAGGGAAGTTTATGAGACCCCTGAACCA 192
 Qy 188 TGAATTTTCTATTATTAAGAGCCAGAACCGAGTCAATTAATTTTCACTTTAAGGCCA 247
 Db 193 TGAATTTTCTATTATTAAGAGCCAGAACCGAGTCAATTAATTTTCACTTTAAGGCCA 252
 Qy 248 ATCTCTGCTGTGACTTTTGAATTAATCTGAGGAGACAGAACATATTTTGTATAGAC 307
 Db 253 ATCTCTGCTGTGACTTTTGAATTAATCTGAGGAGACAGAACATATTTTGTATAGAC 312
 Qy 308 GGAAGGACCTCTGATATTAACAAGAGCCCTTGAACAGGGAACAAGTCTACTCAATC 367
 Db 313 GGAAGGACCTCTGATATTAACAAGAGCCCTTGAACAGGGAACAAGTCTACTCAATC 372
 Qy 368 TCCAGGTTGAGCCCTGAGACGCTAATGGAATTAATGAGAGGCTCAGTCCCTATACCA 427
 Db 373 TCCAGGTTGAGCCCTGAGACGCTAATGGAATTAATGAGAGGCTCAGTCCCTATACCA 432
 Qy 428 TAGAAGTGAAGACATTAACGACATGACCAAGTTTCTCAGTCAAGTACAGAGGCT 487
 Db 433 TAGAAGTGAAGACATTAACGACATGACCAAGTTTCTCAGTCAAGTACAGAGGCT 492
 Qy 488 CAGTAAGGCAAGACTCTGCGCCAGAGAAAGCCCTTCTGTATGTCAAGCAAGACCTGG 547
 Db 493 CAGTAAGGCAAGACTCTGCGCCAGAGAAAGCCCTTCTGTATGTCAAGCAAGACCTGG 552
 Qy 548 ATGATCCGGGCACTCCCAATGCGAGCTTATTAACAGATGTGATCAGACTTCCATGA 607

Db	553	ATGATCCGGCCACTCCCAATGGCCAGCTTTATTACAGATTGTATCCAGCTTCCCATG	612
Qy	608	TCAACAAATGTCATGTACTTTTCAGATCAACAAACAAACGGAGCCATCTCTCTTAACCCGAG	667
Db	613	TCAAACAATGTCATGTACTTTTCAGATCAACAAACAAACGGAGCCATCTCTTAACCCGAG	672
Qy	668	AGGAGTCTCAGGAATTTGAATCTCTGCTAAGAAATCCTTCTATATCTGGTGAATCTCAGTGA	727
Db	673	AGGAGTCTCAGGAATTTGAATCTCTGCTAAGAAATCCTTCTATATCTGGTGAATCTCAGTGA	732
Qy	728	AGCAATGGGAGGCCAGAGTGAAGATTCCTTCAGTATACCAACATCTGTGATATCATAG	787
Db	723	AGCAATGGGAGGCCAGAGTGAAGATTCCTTCAGTATACCAACATCTGTGATATCATAG	792
Qy	788	TGACAGAGATATTTTGAAAGCACCACAAACCTGTGAGATGGTGGAAATCAACTGATC	847
Db	793	TGACAGAGATATTTTGAAAGCACCACAAACCTGTGAGATGGTGGAAATCAACTGATC	852
Qy	848	CTCACCCCATCAAAATCACTCAGTGCGGTGGAATGATCCGGTGCACATATTCCTTAG	907
Db	853	CTCACCCCATCAAAATCACTCAGTGCGGTGGAATGATCCGGTGCACATATTCCTTAG	912
Qy	908	TTGACAAAGAGAGGTGCGCAGAAATCCCATTTTCAATTTGACAGGAAGAGATATTTACG	967
Db	913	TTGACAAAGAGAGGTGCGCAGAAATCCCATTTTCAATTTGACAGGAAGAGATATTTACG	972
Qy	968	TGACTCAGCCCTTGAGCCGAGAAAGAAAGATGCATATGTTTTTATCAGTTGCCAAG	1027
Db	973	TGACTCAGCCCTTGAGCCGAGAAAGAAAGATGCATATGTTTTTATCAGTTGCCAAG	1032
Qy	1028	ATGAGTACGGAAACCACTTTCTATATCCGCTGGAAATTCATGTAAAGTTAAAGATTTA	1087
Db	1033	ATGAGTACGGAAACCACTTTCTATATCCGCTGGAAATTCATGTAAAGTTAAAGATTTA	1092
Qy	1088	ATGATTAATCCACCTACATGTCCGTCACAGTAACCGTATTTGAGTCCAGAGAAATGAC	1147
Db	1093	ATGATTAATCCACCTACATGTCCGTCACAGTAACCGTATTTGAGTCCAGAGAAATGAC	1152
Qy	1148	GACTGGGTACAGTATCGGGACCCCTTACTGCAATGACAGGATGTAAGAAATATCTGCCA	1207
Db	1153	GACTGGGTACAGTATCGGGACCCCTTACTGCAATGACAGGATGTAAGAAATATCTGCCA	1212
Qy	1208	ACAGTTTTCTTAACTACAGATTTGTGAGCAAACTCCCAAACTTCCATGATGGAATCT	1267
Db	1213	ACAGTTTTCTTAACTACAGATTTGTGAGCAAACTCCCAAACTTCCATGATGGAATCT	1272
Qy	1268	TCTTAATCCAAACCTATGCTGGAAATGTTACAGTTAAGCTTAACAGTCCCTTGAAGACAG	1327
Db	1273	TCTTAATCCAAACCTATGCTGGAAATGTTACAGTTAAGCTTAACAGTCCCTTGAAGACAG	1332
Qy	1328	ATACTCCTCAGTACAACTTAAAGATAGAGTGTCTGACAAAGATTTCAAGACCCCTTGT	1387
Db	1333	ATACTCCTCAGTACAACTTAAAGATAGAGTGTCTGACAAAGATTTCAAGACCCCTTGT	1392
Qy	1388	TTGTGCAATCAACGTTATATGATATCAATGATCGAGATCCCATCTTTGAAAATCAAGTT	1447
Db	1393	TTGTGCAATCAACGTTATATGATATCAATGATCGAGATCCCATCTTTGAAAATCAAGTT	1452
Qy	1448	ATGGAACCTGATCTTTGCTGTAAGACAAACATTTGGGTCCACCATCTTAAACATCCAG	1507
Db	1453	ATGGAACCTGATCTTTGCTGTAAGACAAACATTTGGGTCCACCATCTTAAACATCCAG	1512
Qy	1508	CCACTGATGCTGATGAGCCATTTTACTGAGAGTTCCTAAATCTGTATCATATCTTAAG	1567
Db	1513	CCACTGATGCTGATGAGCCATTTTACTGAGAGTTCCTAAATCTGTATCATATCTTAAG	1572
Qy	1568	GAGACAGTGAAGGAGCGCTGGGGGTGACACAGATCCCATACCAACCGGATATGTCA	1627
Db	1573	GAGACAGTGAAGGAGCGCTGGGGGTGACACAGATCCCATACCAACCGGATATGTCA	1632
Qy	1628	TAAATTAAGACCTCTTGAATTTTGAAGACAGCTGTTCCAAATGTTGTTCAAAGAG	1687

Db	1633	TAATTAAAAAGCCTCTTGATTTTGAACAGACGCTGTTTCCACATTGTGTCAAACAG	1632
Qy	1688	AAAAATCCTAGACCTCTAGTGTGTTGGTGTGAAGTACAATGCAAGTCTCTTTGGCCAAAGTTCA	1747
Db	1693	AAAATCCGACCCCTCTAGTGTGTTGGTGTGAAGTACAATGCAAGTCTCTTTGGCCAAAGTTCA	1752
Qy	1748	CGCTTATTGTGACAGATGTGAATGAACACTCAATTTTCCCAACAGTATTCACAGCA	1807
Db	1753	CGCTTATTGTGACAGATGTGAATGAACACTCAATTTTCCCAACAGTATTCACAGCA	1812
Qy	1808	AAGTCAGGAGAGATGTGATCTATAGGCACTAAAGGGGCAATGTGACCTGCCAAGGATCCAG	1867
Db	1813	AAGTCAGGAGAGATGTGATCTATAGGCACTAAAGGGGCAATGTGACCTGCCAAGGATCCAG	1872
Qy	1868	AAGGCTTGACATTAAGCTATTACCTGAGGGGAGACACAAAGAGTTGGCTTTAAATTGACC	1927
Db	1873	AAGGCTTGACATTAAGCTATTACCTGAGGGGAGACACAAAGAGTTGGCTTTAAATTGACC	1932
Qy	1928	ACGTGACTGGTGAATCTTTTAGTGTGCTCCATTGACAGAGAAAGCCGAAAGTCCATATC	1987
Db	1933	ACGTGACTGGTGAATCTTTTAGTGTGCTCCATTGACAGAGAAAGCCGAAAGTCCATATC	1992
Qy	1988	GGGTACAAAGTGTGGCCACAGAACTAGGGGGGTCTTCCTTAAGCTGTGTCCAGAGTTC	2047
Db	1993	GGGTACAAAGTGTGGCCACAGAACTAGGGGGGTCTTCCTTAAGCTGTGTCCAGAGTTC	2052
Qy	2048	ACCTGATCCTTATGATGTGATGAATGACAACCTCCCAAGGCTAGGCCAAAGACTACACGGGCT	2107
Db	2053	ACCTGATCCTTATGATGTGATGAATGACAACCTCCCAAGGCTAGGCCAAAGACTACACGGGCT	2112
Qy	2108	TGTTCTTCTGCAATCCCTCGAGTGCACCTGGAAGTCTCAATTTTGCAGGCTTACTGATGATG	2167
Db	2113	TGTTCTTCTGCAATCCCTCGAGTGCACCTGGAAGTCTCAATTTTGCAGGCTTACTGATGATG	2172
Qy	2168	ATCGACACTTATTCGGGGTCCCATTTTCAATTTTCCCTGGGAGTGAAGACTTACAA	2227
Db	2173	ATCGACACTTATTCGGGGTCCCATTTTCAATTTTCCCTGGGAGTGAAGACTTACAA	2232
Qy	2228	ACGACTGGGAAGTTCCAAATCAATGATGTAATGCTGATGCGGACTGTCTACACGAGCACAG	2287
Db	2233	ACGACTGGGAAGTTCCAAATCAATGATGTAATGCTGATGCGGACTGTCTACACGAGCACAG	2292
Qy	2288	ACTTTGAGAGAGGGGCGTATGTGTCTTGTATCCGCATCAATGATGGGGGTCCGACACCT	2347
Db	2293	AGTTTGAGAGAGGGGAGTATGTGTCTTGTATCCGCATCAATGATGGGGGTCCGACACCT	2352
Qy	2348	TGGAAGGCAATTTGTTCTTTACCACTTACATTTCTGCAATGTGTGTGGAAGAAAGTTGTTCC	2407
Db	2353	TGGAAGGCAATTTGTTCTTTACCACTTACATTTCTGCAATGTGTGTGGAAGAAAGTTGTTCC	2412
Qy	2408	GGCAGAGAGGTCACCAAGCTGGGATACCACTGNGGGCATGGGAGTTGGATPACTGTGA	2467
Db	2413	GGCAGAGAGGTCACCAAGCTGGGATACCACTGNGGGCATGGGAGTTGGATPACTGTGA	2472
Qy	2468	CCACCCCTTCTGTGATTTGTATTAATTTTGTAGCAATGTGTGTTATCCGATTAAGAAGATA	2527
Db	2473	CCACCCCTTCTGTGATTTGTATTAATTTTGTAGCAATGTGTGTTATCCGATTAAGAAGATA	2532
Qy	2528	AAGCCAAAGATTAATGTGTGAAGTGTCAAGCACTGAAAGTCAAACTCTGTGAAGACTGAA	2587
Db	2533	AAGCCAAAGATTAATGTGTGAAGTGTCTCAAGCACTGTGAAGTCAAACTCTGTGAAGACTGAA	2592
Qy	2588	TTTGAAGAAAGATTTGATTTATATATATAGCAAGTGGCAATTTTGAGCAACAAACATCTCATC	2647
Db	2593	TTTGAAGAAAGATTTGATTTATATATATAGCAAGTGGCAATTTTGAGCAACAAACATCTCATC	2652
Qy	2648	CTATTACTTTTCACTAACGTGATTAATAATTTTAAAACAGATATTTCCCTCTGTCTCT	2707
Db	2653	CTATTACTTTTCACTAACGTGATTAATAATTTTAAAACAGATATTTCCCTCTGTCTCT	2712
Qy	2708	TAAATTTTGGTAAATATTTCTTTTTTGAAGGTGAGTCTTGCTGTGTGCCACAGCTGAG	2767
Db	2713	TAAATTTTGGTAAATATTTCTTTTTTGAAGGTGAGTCTTGCTGTGTGCCACAGCTGAG	2772

QY 2768 TACAGTGTGTGATCCAGCTCACTGCAACTCCGCTCCTGGGTTTCAATGATTCCTC 2827
DB 2773 TACAGTGTGTGATCCAGCTCACTGCAACTCCGCTCCTGGGTTTCAATGATTCCTC 2832
QY 2828 GCCTCAGCTTCCCTAAGTGTGATGTTTACAGGACCCACCAATGCCAGTAATTTT 2887
DB 2833 GCCTCAGCTTCCCTAAGTGTGATGTTTACAGGACCCACCAATGCCAGTAATTTT 2892
QY 2888 GATATTTTAAATAGAGAGCGGGGTTTCCGCAATTTGGCCAGGCTGTCTTGAATCCTCGA 2947
DB 2893 GATATTTTAAATAGAGAGCGGGGTTTCCGCAATTTGGCCAGGCTGTCTTGAATCCTCGA 2952
QY 2948 CAAGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3007
DB 2953 CAAGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3012
QY 3008 AGATATTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3067
DB 3013 AGATATTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3072
QY 3068 CTCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3127
DB 3073 CTCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3132
QY 3128 TATTCGTATATTTTCTTATCAAGAGATATATAGTGTGTCTCATAGAACTGCT 3187
DB 3133 TATTCGTATATTTTCTTATCAAGAGATATATAGTGTGTCTCATAGAACTGCT 3192
QY 3188 GATTCATTTTATTTTCTGATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3247
DB 3193 GATTCATTTTATTTTCTGATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3252
QY 3248 GATTCATTTTATTTTCAAAATTTTGTCAAGAAAGAAAGTGAAGATCTGAGAAAT 3307
DB 3253 GATTCATTTTATTTTCAAAATTTTGTCAAGAAAGAAAGTGAAGATCTGAGAAAT 3312
QY 3308 AATTAATTAAGAGAGAGCTTTTGC 3333
DB 3313 AATTAATTAAGAGAGAGCTTTTGC 3338

RESULT 8
AX411163 3654 bp DNA linear PAT 14-JUN-2002
LOCUS AX411163
DEFINITION Sequence 3810 from Patent WO0229103.
ACCESSION AX411163
VERSION AX411163.1 GI:21443868
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS Alvarez, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 3810 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
source Location/Qualifiers
1..3654
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. X83228"

BASE COUNT 1052 a 804 c 777 g 1021 t
ORIGIN

Query Match 99.1%; Score 3314.8; DB 6; Length 3654;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3319; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 8 GTCTGCACTGTAATGGAAGAAAGACTTTTAACCACTTTTGTGACTTACAGAAAG 67
DB 13 GTCTGCACTGTAATGGAAGAAAGACTTTTAACCACTTTTGTGACTTACAGAAAG 72

QY 68 GAATTTGAATTAAGAAAATATGATPACTTCAAGGCCATCTTCACTCCTGTGCTTTA 127
DB 73 GAATTTGAATTAAGAAAATATGATPACTTCAAGGCCATCTTCACTCCTGTGCTTTA 132
QY 128 TGTCTTATTTGGCAACTGTGATATGCGCAAGAGGGAAGTTTATGTGACCCCTGAAACCA 187
DB 133 TGTCTTATTTGGCAACTGTGATATGCGCAAGAGGGAAGTTTATGTGACCCCTGAAACCA 192
QY 188 TGACATTTTCTATTTATGAAGCCAAAGAACCGATCAATTTATCTCCAGTTTAAGGCA 247
DB 193 TGACATTTTCTATTTATGAAGCCAAAGAACCGATCAATTTATCTCCAGTTTAAGGCA 252
QY 248 ATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 307
DB 253 ATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 312
QY 308 GGGAGGGAATCTGTATTAACAAGACCTTGGACAGGGAACAGATCTTACTCAATC 367
DB 313 GGGAGGGAATCTGTATTAACAAGACCTTGGACAGGGAACAGATCTTACTCAATC 372
QY 368 TCGAGGTGCAAGCCCTGAGAGCTTAATGGAATTTATGTAAGGTCAGTCCCATACCA 427
DB 373 TCGAGGTGCAAGCCCTGAGAGCTTAATGGAATTTATGTAAGGTCAGTCCCATACCA 432
QY 428 TAGAAGTGAAGGACATCAACGACATCGAACCCAGTTTCTCCAGTCAAAAGTACGAAGCT 487
DB 433 TAGAAGTGAAGGACATCAACGACATCGAACCCAGTTTCTCCAGTCAAAAGTACGAAGCT 492
QY 488 CAGTAAGGAGAACTCTGCGCCAGGAAGACCTTTCTTGTATGTCAATGCCACAGACTGG 547
DB 493 CAGTAAGGAGAACTCTGCGCCAGGAAGACCTTTCTTGTATGTCAATGCCACAGACTGG 552
QY 548 ATGATCCGCGCACTCCCAATGCGCAGCTTATTAACAAGTGTGATCCAGCTTCCATGA 607
DB 553 ATGATCCGCGCACTCCCAATGCGCAGCTTATTAACAAGTGTGATCCAGCTTCCATGA 612
QY 608 TCAACAAATGTCATGATCTTTCAGATCAACAAACGAGAGCATCTCTTACCCGAG 667
DB 613 TCAACAAATGTCATGATCTTTCAGATCAACAAACGAGAGCATCTCTTACCCGAG 672
QY 668 AGGATCTCAGAAATTAATCTGCTAAGATCTCTTATTAATCTGCTGATCTCAGTGA 727
DB 673 AGGATCTCAGAAATTAATCTGCTAAGATCTCTTATTAATCTGCTGATCTCAGTGA 732
QY 728 AGGATCTCAGAAATTAATCTGCTAAGATCTCTTATTAATCTGCTGATCTCAGTGA 787
DB 733 AGGATCTCAGAAATTAATCTGCTAAGATCTCTTATTAATCTGCTGATCTCAGTGA 792
QY 788 TGACAGAGAAATTTTGAAGAACCAACAAACCTGTGAGATGTGGAACCACTCACTGATC 847
DB 793 TGACAGAGAAATTTTGAAGAACCAACAAACCTGTGAGATGTGGAACCACTCACTGATC 852
QY 848 CTCACCCCATCAAAATCACTCAGGTGCGGTGAATATCCCGGTGCACAATATTCCTTAA 907
DB 853 CTCACCCCATCAAAATCACTCAGGTGCGGTGAATATCCCGGTGCACAATATTCCTTAA 912
QY 908 TTGACAAAGAGAGCTGCGCAAGTTTCCATTTTCAATTTGACCGAAGAGATATTAGC 967
DB 913 TTGACAAAGAGAGCTGCGCAAGTTTCCATTTTCAATTTGACCGAAGAGATATTAGC 972
QY 968 TGACTCAGCCCTTGGACCGGAGAAAGAGATCATATGTTTATTTATGCAAGTCAAGG 1027
DB 973 TGACTCAGCCCTTGGACCGGAGAAAGAGATCATATGTTTATTTATGCAAGTCAAGG 1032
QY 1028 ATGATTAATCGAATAATCACTTTCATATCCGTGGAATTCATGTAATAAGATATTA 1087
DB 1033 ATGATTAATCGAATAATCACTTTCATATCCGTGGAATTCATGTAATAAGATATTA 1092
QY 1088 ATGATTAATCGAATAATCACTTTCATATCCGTGGAATTCATGTAATAAGATATTA 1147
DB 1093 ATGATTAATCGAATAATCACTTTCATATCCGTGGAATTCATGTAATAAGATATTA 1152

QY 1148 GACTGGTAACAGTATGGGACCCCTTACTGACATGACAGGGATGAGAAAAATCTGCC 1207
DB 1153 GACTGGGTAAACAGTATGGGACCCCTTACTGACATGACAGGGATGAGAAAAATCTGCC 1212
QY 1208 ACAGTTTCTTAACTACAGAGTTGTGGAGCAAACTCCCAAACCTCCATGGATGACCTCT 1267
DB 1213 ACAGTTTCTTAACTACAGAGTTGTGGAGCAAACTCCCAAACCTCCATGGATGACCTCT 1272
QY 1268 TCCCTAATCCAAACCTATGCTGGAATGTATACAGTTAGCTTAAACAGTCCCTTGAAGAAGCAG 1327
DB 1273 TCCCTAATCCAAACCTATGCTGGAATGTATACAGTTAGCTTAAACAGTCCCTTGAAGAAGCAG 1332
QY 1328 ATACTCTCAGTACATCTTAACGATAGAGGTGTCTGACAAAGATTTCAAGACCTTTGTT 1387
DB 1333 ATACTCTCAGTACATCTTAACGATAGAGGTGTCTGACAAAGATTTCAAGACCTTTGTT 1392
QY 1388 TTGTGCAAAATCAACGTTATGATATCAATGATCAAGATCCCATCTTTGAAAAATCAAGATT 1447
DB 1393 TTGTGCAAAATCAACGTTATGATATCAATGATCAAGATCCCATCTTTGAAAAATCAAGATT 1452
QY 1448 ATGGAACCTGACTCTTGTGTAAGACAAACATTTGGGTCCACCATCTTAACATCCAGG 1507
DB 1453 ATGGAACCTGACTCTTGTGTAAGACAAACATTTGGGTCCACCATCTTAACATCCAGG 1512
QY 1508 CCACTGATGCTGATGAGCCATTTACTGGAGTTCTAAATTTCTGTATCATATCAATAAGG 1567
DB 1513 CCACTGATGCTGATGAGCCATTTACTGGAGTTCTAAATTTCTGTATCATATCAATAAGG 1572
QY 1568 GAGACAGTGAAGGACGCTGGGGGTGACACAGATCCCATACCAACACCGGATATGTCA 1627
DB 1573 GAGACAGTGAAGGACGCTGGGGGTGACACAGATCCCATACCAACACCGGATATGTCA 1632
QY 1628 TAATTAATAAAGCCTCTGATTTTGAAGACAGACGCTGTTCCAACTATGTTCCAAAGCAG 1687
DB 1633 TAATTAATAAAGCCTCTGATTTTGAAGACAGACGCTGTTCCAACTATGTTCCAAAGCAG 1692
QY 1688 AAAATCTGAGCCTCTAGTGTGTGTGTGAAGTACAAATGCAAGTTCTTTTGGCAAGTTCA 1747
DB 1693 AAAATCTGAGCCTCTAGTGTGTGTGTGAAGTACAAATGCAAGTTCTTTTGGCAAGTTCA 1752
QY 1748 CGCTTATGTGACAGATGTAAGTAAGCACTCAATTTTCCCAACAGTATTCMAAGCA 1807
DB 1753 CGCTTATGTGACAGATGTAAGTAAGCACTCAATTTTCCCAACAGTATTCMAAGCA 1812
QY 1808 AAGTCACTGAGATGTAGTATAGCACTAAAGTGGCAATGTGACTGCCAAGGATCCAG 1867
DB 1813 AAGTCACTGAGATGTAGTATAGCACTAAAGTGGCAATGTGACTGCCAAGGATCCAG 1872
QY 1868 AAGTCTGACATTAAGCTATTACTGAGGGGAGACAAAGAGTTGGCTTAAATTTGACC 1927
DB 1873 AAGTCTGACATTAAGCTATTACTGAGGGGAGACAAAGAGTTGGCTTAAATTTGACC 1932
QY 1928 AAGTGACTGAGATTTTATAGTGTGCTCAATTTGACAGAGAAACCGGAAATGCCATATC 1987
DB 1933 AAGTGACTGAGATTTTATAGTGTGCTCAATTTGACAGAGAAACCGGAAATGCCATATC 1992
QY 1988 GGGTAACAAGTGTGGGACCAAGAGTAGGGGGTCTTCTTAAGTCTGTGTCAGAGTTCC 2047
DB 1993 GGGTAACAAGTGTGGGACCAAGAGTAGGGGGTCTTCTTAAGTCTGTGTCAGAGTTCC 2052
QY 2048 ACCTGATCTTATGATGTAATGACAAACCTCCAGAGCTAGCCAAAGACTACAGGGCT 2107
DB 2053 ACCTGATCTTATGATGTAATGACAAACCTCCAGAGCTAGCCAAAGACTACAGGGCT 2112
QY 2108 TGTTCCTTGTGCAATCCCTCAGTGCACCTGGAAGTCTCATTTTGCAGGCTTACTGATGATG 2167
DB 2113 TGTTCCTTGTGCAATCCCTCAGTGCACCTGGAAGTCTCATTTTGCAGGCTTACTGATGATG 2172
QY 2168 ATCAGCACTTATTTGGGGTCCCATTTTTCATTTTCCCTGGAGGTGGAAGCTTACAA 2227
DB 2173 ATCAGCACTTATTTGGGGTCCCATTTTTCATTTTCCCTGGAGGTGGAAGCTTACAA 2232
QY 2228 ACGACTGGGAAGTTTCCAAATCAATGTACTCATGCCCAGTGTCTTACAGGACACAG 2287

DB 2233 ACGACTGGGAAGTTTCCAAATCAATGTACTCATGCCCACATGTCTTACAGGACACAG 2292
QY 2288 ACTTTGAGAGAGGGGGTATGTGCTGTGATCCGCAATCAATGATGGGGGTGGGCAACCT 2347
DB 2293 AGTTTGGAGAGAGGATATGTGTCTTGATCCGCAATCAATGATGGGGGTGGGCAACCT 2352
QY 2348 TGGAAAGCATTTGTTCTTTTACAGATTACATCTGCAAGTTGTGGAAGAAAGTTGTTCC 2407
DB 2353 TGGAAAGCATTTGTTCTTTTACAGATTACATCTGCAAGTTGTGGAAGAAAGTTGTTCC 2412
QY 2408 GGCACAGAGGTCAACAGACTGGGATACCACTGTGGGCATGCGAGTTGGTATCTGCTGA 2467
DB 2413 GGCACAGAGGTCAACAGACTGGGATACCACTGTGGGCATGCGAGTTGGTATCTGCTGA 2472
QY 2468 CCAACCTTCTGGGATTTGGATATTTTGAAGTGTGAGTGTGTTATCCGATTAAGAAGATA 2527
DB 2473 CCAACCTTCTGGGATTTGGATATTTTGAAGTGTGAGTGTGTTATCCGATTAAGAAGATA 2532
QY 2528 AAGCAAAAGATTAATGTGAAAGTCTCAAGCATCTGAAGTCAAACTCTGAGAAAGTGA 2587
DB 2533 AAGCAAAAGTAAATGTGAAAGTCTCAAGCATCTGAAGTCAAACTCTGAGAAAGTGA 2592
QY 2588 TTTGAAAAGATGTTGAAATTTTATATAGCAAGTGTATTTAGCAACCAATCTTCATC 2647
DB 2593 TTTGAAAAGATGTTGAAATTTTATATAGCAAGTGTATTTAGCAACCAATCTTCATC 2652
QY 2648 CTATTACTTTTCACTCAAGTGTATTAATTTTAAACAGATTTTCCCTCTGTCTCT 2707
DB 2653 CTATTACTTTTCACTCAAGTGTATTAATTTTAAACAGATTTTCCCTCTGTCTCT 2712
QY 2708 TAATATTTGCTAAATATTTCTTTTGAAGTGAAGTCTGTGCTGTGCGCCAGGCTGAG 2767
DB 2713 TAATATTTGCTAAATATTTCTTTTGAAGTGAAGTCTGTGCTGTGCGCCAGGCTGAG 2772
QY 2768 TACAGTGTGTGATCCCAAGCTCACTGCAACCTCCGCTCTGTGGTTCAATGATCTCT 2827
DB 2773 TACAGTGTGTGATCCCAAGCTCACTGCAACCTCCGCTCTGTGGTTCAATGATCTCT 2832
QY 2828 GCCTGAGCTTCCCAATGATGAGTGTGTTTACAGAGACCCACACATGCCCAAGCTAATTTT 2887
DB 2833 GCCTGAGCTTCCCAATGATGAGTGTGTTTACAGAGACCCACACATGCCCAAGCTAATTTT 2892
QY 2888 GTATTTTAAATAGAGAGGGGGTTCGCAATTTGGCCAGGCTGCTTGAACCTCTGACGT 2947
DB 2893 GTATTTTAAATAGAGAGGGGGTTCGCAATTTGGCCAGGCTGCTTGAACCTCTGACGT 2952
QY 2948 CAAATGATCTGCTGCTTGTGCTCCCAATACAGGATGAAACCACTGCAACCACTTAT 3007
DB 2953 CAAATGATCTGCTGCTTGTGCTCCCAATACAGGATGAAACCACTGCAACCACTTAT 3012
QY 3008 AGATATTTCAATGTGCTATAGACATTAAGAGATTTTCAATTTTCCATGACATTTTCC 3067
DB 3013 AGATATTTCAATGTGCTATAGACATTAAGAGATTTTCAATTTTCCATGACATTTTCC 3072
QY 3068 CTCTGAAATAGGCTTGAAGTGTGTTTCCCTTTGGGGCAAGACGACTTAA 3127
DB 3073 CTCTGAAATAGGCTTGAAGTGTGTTTCCCTTTGGGGCAAGACGACTTAA 3132
QY 3128 TATTTGTAATTTTCTTTATCAAGAGATATATCAAGTGTGCTCATAGAACTGCT 3187
DB 3133 TATTTGTAATTTTCTTTATCAAGAGATATATCAAGTGTGCTCATAGAACTGCT 3192
QY 3188 GGAATTCATTTATGTTTCTTATGATTCATCTGTGTCCCTTCACTCTTGG 3247
DB 3193 GGAATTCATTTATGTTTCTTATGATTCATCTGTGTCCCTTCACTCTTGG 3252
QY 3248 GTATTCACATGAAATTTCAAAATTTGTGAGAGAAAGAAAGAGAGATCCAGAAAT 3307
DB 3253 GTATTCACATGAAATTTCAAAATTTGTGAGAGAAAGAAAGAGAGATCCAGAAAT 3312
QY 3308 AAATAAATAAAGAACAGCCTTTTGC 3333

Db 3313 AATTAATAAGACGCTTTTC 3338

RESULT 9
HSTRNALICA 3654 bp mRNA linear PRI 01-JUN-1995
LOCUS H.sapiens mRNA for LI-cadherin.
DEFINITION X83228
ACCESSION X83228.1 GI:854174
VERSION 1
KEYWORDS LI-cadherin.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3654)
Boettlinger, A., Kreitz, B., Fieger, C., Dlouhy, B., Berndorff, D.,
Goessner, R. and Tauber, R.
Molecular cloning of human LI-cadherin: evidence for a novel type of
cadherin within the cadherin superfamily
TITLE Unpublished
JOURNAL 2 (bases 1 to 3654)
REFERENCE Boettlinger, A.M.
AUTHORS Direct Submission
TITLE Submitted (05-DEC-1994) A.M. Boettlinger, Institut fuer Klinische
JOURNAL Chemie & Bioch., Universitaetsklinikum Rudolf Virchow, Freie
Universitaet Berlin, Spandauer Damm 130, 14050 Berlin, FRG
COMMENT Sequence overlapping with that under the accession number U07969.
FEATURES
source location/Qualifiers
1..3654
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="uni-ZAP TM vector"
93..2591
/function="cell adhesion molecule"
/codon_start=1
/product="LI-cadherin"
/protein_id="CAA58231.1"
/db_xref="GI:854175"
/db_xref="SPTREMBL:Q15336"
/translation="MILQAHLSICLMLYANGYGEKFGSEPLKMTSYIEGQEP
SOLIFQKPNPAPATFELTGTETDNI FVIEREGSLYNRALDRERSTHNIQVALDNI
GILVEGVPITIKVDINDRPTFLOSREGSVYONRPGPELYVATLDLPATPN
GOLYVOIVLOPMNNMYPOINKKTAISLREGSOELNKAAPS,NLIVSYVDMGQ
QSEMSRSDTTSVDIIVTENTINKKPIEMVENSDBPHITIKTOVRMDPAQSLVXK
EKPRFPFSIDQSDIYVTPDLEEKDAYFYVAADRYGKPLSYLEIHHVKYKDN
DNPTCSPTVFEVQENRIGNSIGTLTAHDEENTANSFLNRYVEQTPKLPMDG
LFLIOTVAGMLQAKOSLKKQDTPQVNLTEVSDKDKFLCFQVINVIDINDPPIE
KSDYGNLTLEADTNI GSTILTI QATDADEPTGSKILYHIKDGSEGRIGVDTPHT
NTGVIIKKPLDETAIVSNIVFEKPNBPVFGVKNASFAKFTLIVDVNEAPQF
SQHYRQAKVSEDAVIGTKVGNVTKKDEGLDISLGGDRGMLKIDHVFGEISVAP
LDRFAGSPYRQVAVATEVGGSSLSVSEFHLILMDVNDNPRKLAKDTYGLFCHPLSA
PGSLIFPATDDQHLFRGPHTFSLGSSQLQNDWEVKINGSTHARLSRTRFEERYEY
VVLIRINDGGRPLLEGIIVSLVPTFCSCVSGECPRPACHQIGIPIVGNVAVGILLTLTLLV
IGTLIAVFEIRIKDKGNDVESAQASEVFKLRS"

BASE COUNT 1052 a 804 c 777 g 1021 t

ORIGIN

Query Match 99.1%; Score 3314.8; DB 9; Length 3654;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3319; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 8 GTCTCGACCTGTAATGAGAGAAAGACCTTTAAACCCATTTTGTGACTTACAGAAAG 67
13 GTCTCGACCTGTAATGAGAGAAAGACCTTTAAACCCATTTTGTGACTTACAGAAAG 72

Qy 68 GAATTTGAATAAGAAACCTATGATCTTCAAGGCCCATCTTCACTCCCTGTGTCTTCTTA 127
132 GTCTCGACCTGTAATGAGAGAAAGACCTTTAAACCCATTTTGTGACTTACAGAAAG 132

Qy 128 TGCTTTATTTGGCAACTGATATGCGCAAGAGGGAGATTGAGACCCCTGAAACCA 187
133 TGCTTTATTTGGCAACTGATATGCGCAAGAGGGAGATTGAGACCCCTGAAACCA 192

Qy 188 TGACATTTTCTATTTATGAAAGCCAGAAACCGAGTCAAATTTATTCAGTTTAAGGCCA 247
193 TGACATTTTCTATTTATGAAAGCCAGAAACCGAGTCAAATTTATTCAGTTTAAGGCCA 252

Qy 248 ATCTCTCTGTGTGACTTTTGAACCTAACCTGGGGAGACAGAACATATTTTGTATGAAAC 307
253 ATCTCTCTGTGTGACTTTTGAACCTAACCTGGGGAGAGAGAACATATTTTGTATGAAAC 312

Qy 308 GGGAGGACCTTGTATTTTCAACAGAGCCCTTGGACAGGGGAAACAGATCTACTCAATC 367
313 GGGAGGACCTTGTATTTTCAACAGAGCCCTTGGACAGGGGAAACAGATCTACTCAATC 372

Qy 368 TCAGAGTTGAGCCCTGAGACGCTAATGAAATTATGAGAGGCTCCATGACCA 427
373 TCAGAGTTGAGCCCTGAGACGCTAATGAAATTATGAGAGGCTCCATGACCA 432

Qy 428 TAGAAGTGAAGACATCAACGACATGACCCAGCTTCTCCAGTCAAAAGTACGAAGGCT 487
433 TAAAGTGAAGACATCAACGACATGACCCAGCTTCTCCAGTCAAAAGTACGAAGGCT 492

Qy 488 CAGTAAGGAGAACTCTGCGCCAGAAAGCCCTTGTATGTCATGACAGACAGCCTGG 547
493 CAGTAAGGAGAACTCTGCGCCAGAAAGCCCTTGTATGTCATGACAGACAGCCTGG 552

Qy 548 ATGATCCGGCCACTCCCAATGCGACGCTTATTAACAGATTCATCCAGCTTCCATGA 607
553 ATGATCCGGCCACTCCCAATGCGACGCTTATTAACAGATTCATCCAGCTTCCATGA 612

Qy 608 TCAACAATGTCATGATCTTCAATCAACAAACAAACGGAGCCATCTCTTAACCCGAG 667
613 TCAACAATGTCATGATCTTCAATCAACAAACAAACGGAGCCATCTCTTAACCCGAG 672

Qy 668 AGGATCTCGAGAAATTAATCCCTGCTAAGAAATCCCTTATATCTGGTATCTCAGTGA 727
673 AGGATCTCGAGAAATTAATCCCTGCTAAGAAATCCCTTATATCTGGTATCTCAGTGA 732

Qy 728 AGGACATGGAGGCGCAGAGTGAAGAAATTCCTTCAGTATACCAATCTGTGATATCATAG 787
733 AGGACATGGAGGCGCAGAGTGAAGAAATTCCTTCAGTATACCAATCTGTGATATCATAG 792

Qy 788 TGAAGAGAAATTTTGAAGAAACCAACCTGTGAGATGTGAGAAACCTCAATGATC 847
793 TGAAGAGAAATTTTGAAGAAACCAACCTGTGAGATGTGAGAAACCTCAATGATC 852

Qy 848 CTGACCCCATCAAAATCACTCAGGTGGGGGAAATATCCCGGACCAATATTCCTTAG 907
853 CTGACCCCATCAAAATCACTCAGGTGGGGGAAATATCCCGGACCAATATTCCTTAG 912

Qy 908 TTGACAAAGAGAGCTGCCAAGATTCCTCATTTTCAATTGACAGAGAGATATTACG 967
913 TTGACAAAGAGAGCTGCCAAGATTCCTCATTTTCAATTGACAGAGAGATATTACG 972

Qy 968 TGACTGAGCCCTTGGACCGAGAAAGAAAGATGATATGTTTTTATGCACTTGCAAAAG 1027
973 TGACTGAGCCCTTGGACCGAGAAAGAAAGATGATATGTTTTTATGCACTTGCAAAAG 1032

Qy 1028 ATGAGTACGGAAGAAACCACTTTCATATCCGCTGGAATTTCAATGAAAGTAAAGATTTA 1087
1033 ATGAGTACGGAAGAAACCACTTTCATATCCGCTGGAATTTCAATGAAAGTAAAGATTTA 1092

Qy 1088 ATGATATCCACCTACATGTCCTGTCACACAGTAACCGATTTTGAAGTCCAGAGATGAAAC 1147
1093 ATGATATCCACCTACATGTCCTGTCACACAGTAACCGATTTTGAAGTCCAGAGATGAAAC 1152

Qy 1148 GACTGGGTAAACAGTATGAGGACCTTTATCTGACATGACAGGATGAAAGAAATATCTGCCA 1207
1153 GACTGGGTAAACAGTATGAGGACCTTTATCTGACATGACAGGATGAAAGAAATATCTGCCA 1212

Qy 1208 ACACTTTTCTAAATCTACAGATTTGTGAGAGAAATCCCAAACTTCCATGAGATGAACTCT 1267
1213 ACACTTTTCTAAATCTACAGATTTGTGAGAGAAATCCCAAACTTCCATGAGATGAACTCT 1272

Qy 1268 TCCTAATCCAAACCTATGCTGAAATGTATACGTTAGCTAAACAGTCTTGAAGAGCAG 1327

KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2499)
AUTHORS Danczig, A.H., Hoskins, J.A., and Skatrud, P.L.
TITLE Mammalian influx peptide transporter
JOURNAL Patent: US 562085-A 2 15-APR-1997;
FEATURES Location/Qualifiers
Source 1. 2499
BASE COUNT 728 a 559 c 570 g 642 t
ORIGIN
Query Match 74.7%; Score 2499; DB 6; Length 2499;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

88 ATGATATTTCAGGCGCCATCTTCACTCCCTGCTGCTTCTTATGCTTATTTGGCAACTGGA 147
1 ATGATATTTCAGGCGCCATCTTCACTCCCTGCTGCTTCTTATGCTTATTTGGCAACTGGA 60
148 TATGCGCAAGAGGGGAAGTTAGTGAACCCCTGAAACCCATGACATTTTCTATTTATGAA 207
61 TATGCGCAAGAGGGGAAGTTAGTGAACCCCTGAAACCCATGACATTTTCTATTTATGAA 120
208 GGGCAGAACCCGATCAATTAATTCAGTTTAAAGCCAACTCTCTGCTGACTTTT 267
121 GGGCAGAACCCGATCAATTAATTCAGTTTAAAGCCAACTCTCTGCTGACTTTT 180
268 GAATTAATCTGGGGAGACAGACAAATTTTGTATGAAAGGGGGGAGCTTCTGATTTAC 327
181 GAATTAATCTGGGGAGACAGACAAATTTTGTATGAAAGGGGGGAGCTTCTGATTTAC 240
328 AACAGAGCTTGGACAGGGGAAACAAGATCTACTCAATCTCCAGTTTGCAGCCCTGAC 387
241 AACAGAGCTTGGACAGGGGAAACAAGATCTACTCAATCTCCAGTTTGCAGCCCTGAC 300
388 GCTAATGAAATTAATGATGAGGGGTCCTCACTCCCTATCACCATTAAGATGAGACATCAAC 447
301 GCTAATGAAATTAATGATGAGGGGTCCTCACTCCCTATCACCATTAAGATGAGACATCAAC 360
448 GACAAATGACCCAGCTTCTCCAGTCAAAAGTACGAAGGCTCAGTAAGGCAAGCTCTGCG 507
361 GACAAATGACCCAGCTTCTCCAGTCAAAAGTACGAAGGCTCAGTAAGGCAAGCTCTGCG 420
508 CCAGGAAAGCCCTTCTGTATGTCAATGCCACAGACCTGATGATCCGGCCACTCCCAAT 567
421 CCAGGAAAGCCCTTCTGTATGTCAATGCCACAGACCTGATGATCCGGCCACTCCCAAT 480.
568 GGGCAGCTTTATACAGATTTGTCACTCAAGCTTCCCATGATCAACAATGTCAATGACTTT 627
481 GGGCAGCTTTATACAGATTTGTCACTCAAGCTTCCCATGATCAACAATGTCAATGACTTT 540
628 CAGATCAACAACAACGAGGAGCCATCTCTTTACCGAGAGGGATCTCAGAAATTTGAT 687
541 CAGATCAACAACAACGAGGAGCCATCTCTTTACCGAGAGGGATCTCAGAAATTTGAT 600
688 CCTGTAAGATCTTCTTATATCTGTGATCTTCAAGTGAAGGACATGGAGGCCAGAGT 747
601 CCTGTAAGATCTTCTTATATCTGTGATCTTCAAGTGAAGGACATGGAGGCCAGAGT 660
748 GAGAATTCCTTCAAGTATACCAATCTGTGATATCATATGACAGAGATATTTTGAAA 807
661 GAGAATTCCTTCAAGTATACCAATCTGTGATATCATATGACAGAGATATTTTGAAA 720
808 GACACCAAAACCTGTGAGATGAGTGAACCAATCAATGATCTCCACCCCATCAAAATCTCT 867
721 GACACCAAAACCTGTGAGATGAGTGAACCAATCAATGATCTCCACCCCATCAAAATCTCT 780
868 CAGGTGCGGTGATGATCCGGTGACAAATTTCTTATGTTGACAAAGAGAGCTGGCA 927
781 CAGGTGCGGTGATGATCCGGTGACAAATTTCTTATGTTGACAAAGAGAGCTGGCA 840

928 AGATCCCAATTTTCAATGACCAAGAGAGATATTTAGTGACTCAGCCCTTGACCGA 987
841 AGATCCCAATTTTCAATGACCAAGAGAGATATTTAGTGACTCAGCCCTTGACCGA 900
988 GAAGAAAGAGATGATATGTTTTTTATGCAAGTTGCAAGAGATGATGAGGAAACCACTT 1047
901 GAAGAAAGAGATGATATGTTTTTTATGCAAGTTGCAAGAGATGATGAGGAAACCACTT 960
1048 TCATATCCGTGGAATTTCAATGTAAGTAAAGATTTAAATGATATCCACTACATGT 1107
961 TCATATCCGTGGAATTTCAATGTAAGTAAAGATTTAAATGATATCCACTACATGT 1020
1108 CCGTCAACAGTAAACGATTTTGAAGTCCAGAGATGAACGACTGGGTAAACATATCGGG 1167
1021 CCGTCAACAGTAAACGATTTTGAAGTCCAGAGATGAACGACTGGGTAAACATATCGGG 1080
1168 ACCCTTACTGCAATGACAGGATGAAGAAATATCTGCCAAGTTTTCTTAACTACAGG 1227
1081 ACCCTTACTGCAATGACAGGATGAAGAAATATCTGCCAAGTTTTCTTAACTACAGG 1140
1228 ATTGTGAGCAAACTCCCAACTTTCCGATGATGACTCTTCTTAATCCAACTGATGCT 1287
1141 ATTGTGAGCAAACTCCCAACTTTCCGATGATGACTCTTCTTAATCCAACTGATGCT 1200
1288 GGAATGTTACGTTAGCTAAACAGTCTTGAAGAGAGATATCTCTCAGTACAACTTA 1347
1201 GGAATGTTACGTTAGCTAAACAGTCTTGAAGAGAGATATCTCTCAGTACAACTTA 1260
1348 ACAGTAAAGGTGTCTGACAAAGATTTCAAGACCTTTGTTTGTGCAATCAACGTTATT 1407
1261 ACAGTAAAGGTGTCTGACAAAGATTTCAAGACCTTTGTTTGTGCAATCAACGTTATT 1320
1408 GATATCAATATCAATGCCATCTTTGAAATTCAGATTTATGAAACCTGATGCT 1467
1321 GATATCAATATCAATGCCATCTTTGAAATTCAGATTTATGAAACCTGATGCT 1380
1468 GAAGACCAAAACCTGGGTCACCATCTTAACCATCAGGSCACTGATGATGAGCCA 1527
1381 GAAGACCAAAACCTGGGTCACCATCTTAACCATCAGGSCACTGATGATGAGCCA 1440
1528 TTTACTGGAGTTCTAAATTTCTGTATCATATCAATAAAGGAGACAGTGAAGGCTTGT 1587
1441 TTTACTGGAGTTCTAAATTTCTGTATCATATCAATAAAGGAGACAGTGAAGGCTTGT 1500
1588 GGGGTGACCAAGATCCCATACCAACACCGGATATGTCTAATTAAGGCTCTGAT 1647
1501 GGGGTGACCAAGATCCCATACCAACACCGGATATGTCTAATTAAGGCTCTGAT 1560
1648 TTTGAAACAGAGCTTTTCCAAATTTGTGTTCAAGACAGAAATCTGAGCCTCTAGTG 1707
1561 TTTGAAACAGAGCTTTTCCAAATTTGTGTTCAAGACAGAAATCTGAGCCTCTAGTG 1620
1708 TTTGTGTGAAGTACAAATGCAAGTTCTTTTCCAAAGTTCAAGCTTATTTGACAGATGTG 1767
1621 TTTGTGTGAAGTACAAATGCAAGTTCTTTTCCAAAGTTCAAGCTTATTTGACAGATGTG 1680
1768 AATGAGCACTCAATTTTCCCAACAGTATTCGAAGGAAAGTCAATGAGTGAAGT 1827
1681 AATGAGCACTCAATTTTCCCAACAGTATTCGAAGGAAAGTCAATGAGTGAAGT 1740
1828 AATGAGCACTAAAGGGGCAATGATGATGCCAAGGATCAGAAAGGTCTGACATTAAGCTAT 1887
1741 AATGAGCACTAAAGGGGCAATGATGATGCCAAGGATCAGAAAGGTCTGACATTAAGCTAT 1800
1888 TCACTGAGGGGAGACAAAGAGTTGGCTTAAATTTGACCAAGTACTGATGATCTTT 1947
1801 TCACTGAGGGGAGACAAAGAGTTGGCTTAAATTTGACCAAGTACTGATGATCTTT 1860
1948 AGTGTGCTCATTTGACAGAGAAACCGGAAAGTCAATATGGGGTCAAAAGTGTGGCCACA 2007
1861 AGTGTGCTCATTTGACAGAGAAACCGGAAAGTCAATATGGGGTCAAAAGTGTGGCCACA 1920

Db 1201 GGAATGTTACAGTTAGCTAAAGAGTCTCTGAGAGAGCAAGTATCTCTAGTACAACTTA 1260
 Qy 1348 ACAGTAGAGGTGTCTGACAAAGATTTCAAGACCTTTGTTTGTGCAATACGTTAT 1407
 Db 1261 ACAGTAGAGGTGTCTGACAAAGATTTCAAGACCTTTGTTTGTGCAATACGTTAT 1320
 Qy 1408 GATTCATATGATCGATCCCATCTTTGAAATAACATTAATGAAACCTGACTTGT 1467
 Db 1321 GATTCATATGATCGATCCCATCTTTGAAATAACATTAATGAAACCTGACTTGT 1380
 Qy 1468 GAAGACACAAACATTTGGGTCCACCATCTTAAACATCCAGGCACTGATGATGAGCA 1527
 Db 1381 GAAGACACAAACATTTGGGTCCACCATCTTAAACATCCAGGCACTGATGATGAGCA 1440
 Qy 1528 TTTACTGGAGTTCTAAATTTCTGTATCATATCATTAAGGAGACAGTGAAGGAGCGCTG 1587
 Db 1441 TTTACTGGAGTTCTAAATTTCTGTATCATATCATTAAGGAGAGACAGTGAAGGAGCGCTG 1500
 Qy 1588 GGGGTTCACACAGATCCCATCCACACACCGGATATGTCATTAATAAAGCCCTCTGAT 1647
 Db 1501 GGGGTTCACACAGATCCCATCCACACACCGGATATGTCATTAATAAAGCCCTCTGAT 1560
 Qy 1648 TTTGAAACAGAGCTGTTTCCACATTTGTTTCAAGAGCAAGAAATCCCTGAGCCTCTAGT 1707
 Db 1561 TTTGAAACAGAGCTGTTTCCACATTTGTTTCAAGAGCAAGAAATCCCTGAGCCTCTAGT 1620
 Qy 1708 TTTGTTGTAAGTACATGCAATGTTCTTTTGCCACAGTTCCAGCTTATTTGACAGATGTG 1767
 Db 1621 TTTGTTGTAAGTACATGCAATGTTCTTTTGCCACAGTTCCAGCTTATTTGACAGATGTG 1680
 Qy 1768 AATGACACCTCAATTTTCCCAACAGTATCCAGAGGAAAGTCAAGAGAGTATGCT 1827
 Db 1681 AATGACACCTCAATTTTCCCAACAGTATCCAGAGGAAAGTCAAGAGAGTATGCT 1740
 Qy 1828 AATGACACCTCAATTTTCCCAACAGTATCCAGAGGAGTCTGACATTAAGCTAT 1887
 Db 1741 AATGACACCTCAATTTTCCCAACAGTATCCAGAGGAGTCTGACATTAAGCTAT 1800
 Qy 1888 TCAGTGAAGGGGAGACACAAAGGTTGGCTTAAATGACACAGTGTGAGATCTTT 1947
 Db 1801 TCAGTGAAGGGGAGACACAAAGGTTGGCTTAAATGACACAGTGTGAGATCTTT 1860
 Qy 1948 AGTGTGCTCATTTGACAGAGAGCCGGAAGTCCATATCGGGTACAAAGTGTGAGCA 2007
 Db 1861 AGTGTGCTCATTTGACAGAGAGCCGGAAGTCCATATCGGGTACAAAGTGTGAGCA 1920
 Qy 2008 GAAATGAGGGGGGTCTTCTTAAGCTGTGTGACAGTTCACATGATCTTATGATGTG 2067
 Db 1921 GAAATGAGGGGGGTCTTCTTAAGCTGTGTGACAGTTCACATGATCTTATGATGTG 1980
 Qy 2068 AATGACACCTCCAGGCTACAGAGAGTACACAGGCTTGTCTTGTCCATCCCTC 2127
 Db 1981 AATGACACCTCCAGGCTACAGAGAGTACACAGGCTTGTCTTGTCCATCCCTC 2040
 Qy 2128 AGTGCACCTGGAAGTCTATTTGAGGCTGTGATGATGATGACATTAATTTGAGGGGT 2187
 Db 2041 AGTGCACCTGGAAGTCTATTTGAGGCTGTGATGATGATGACATTAATTTGAGGGGT 2100
 Qy 2188 CCCCATTTTACATTTTCCCTCGGAGTGAAGCTTAAACAAAGCTGGAAGTTTCCAA 2247
 Db 2101 CCCCATTTTACATTTTCCCTCGGAGTGAAGCTTAAACAAAGCTGGAAGTTTCCAA 2160
 Qy 2248 ATCAATGATCTATGCTCCGAGCTGTCTACAGGACACAGACTTTGAGAGAGGGCTAT 2307
 Db 2161 ATCAATGATCTATGCTCCGAGCTGTCTACAGGACACAGACTTTGAGAGAGGGCTAT 2220
 Qy 2308 GTGCTTGAATCCGATCAATGAGGGGGTGGCCACCTTGAAGGATTTGTTCTTA 2367
 Db 2221 GTGCTTGAATCCGATCAATGAGGGGGTGGCCACCTTGAAGGATTTGTTCTTA 2280
 Qy 2368 CCAGTTACATTTCTGAGTGTGTGAGAGGAAGTTGTTTCCGAGCAGAGTACACAGACT 2427

Db 2281 CCAGTTACATTTCTGACAGTTGTGTGAGAGAGTTGTTTCCGACACAGCTCACAGACT 2340
 Qy 2428 GGGATGCCACCTGTGGGAGATGGAGTGTGATATCTGCTGACACCCCTTGTGATGTG 2487
 Db 2341 GGGATGCCACCTGTGGGAGATGGAGTGTGATATCTGCTGACACCCCTTGTGATGTG 2400
 Qy 2488 ATATTTTGAAGCTTGTGTTTATCCGATTAAGAGATTAAGCAAGATTAATGTTGAA 2547
 Db 2401 ATATTTTGAAGCTTGTGTTTATCCGATTAAGAGATTAAGCAAGATTAATGTTGAA 2460
 Qy 2548 AGTGTCAAGCATCTGAGTCAAACTCTGAGAGACTGA 2586
 Db 2461 AGTGTCAAGCATCTGAGTCAAACTCTGAGAGACTGA 2499
 RESULT 12
 RNCADHER
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Rattus norvegicus.
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 3628)
 Berndorff, D., Gessner, R., Kreft, B., Schroy, N., Lajous-Petec, A.M.,
 Loch, N., Reuteler, W., Hortsch, M., and Tauber, R.
 Liver-intestine cadherin: molecular cloning and characterization of
 a novel Ca(2+)-dependent cell adhesion molecule expressed in liver
 and intestine
 J. Cell Biol. 125 (6), 1353-1369 (1994)
 MEDLINE
 PUBMED
 8207063
 2 (bases 1 to 3628)
 Tauber, R.
 Direct Submission
 Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie
 und Biochemie, Universitaet Klinikum Rudolf Virchow, Freie Uni.
 Berlin, Spandauer Damm 130, 14050 Berlin, FRG
 Location/Qualifiers
 1..3628
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /sex="male"
 /tissue_type="liver"
 /clone_lib="lambda ZapII Gene bank, Stratagene Catalog #
 936507"
 /dev_stage="7-8 weeks"
 328..2811
 /codon_start=1
 /product="cadherin"
 /protein_id="CA55631.1"
 /db_xref="GI:505563"
 /db_xref="SWISS-PROT:P55281"
 /translation="MVSQAHFCLTLVITGAYGQGRSGPLKPTPTSPFGQERS
 QIIFQFANPAPVATFELTGETDIFLEKQGLLYHRVLDRETRAVAHLLDLALDSG
 AIVDGVPILIEVDKINDRPTLFQTEBESVQNSRPGKPFMYVNAVLQDLDPATNG
 QLPYQIYIQPKINNNWYFOIDNKGTGALSIPGSGVLDIKNPYNLVVSVDMGQ
 NENSFSDTTSVDTITRENIWKAPBPVIRENLDPHPKIKITVOOMDPGAYSLINKE
 KLPQFPSTIDQENITVYTOPLDREKDSHFPTAADENSKPLAYPEIFVKYDIND
 NPPTCLSQVTFEVEQENEVGSIGIFAHMDNBANNINSILKRLVDTQPKYPSDEL
 FLIDEGKQVQLKRSLLKQDSQVNLIVESDIIDKTLCSLVNVLINDQILPFR
 SDYGSKLSIEDTIAIGSTILIIQATDDEPFTGSKLLYKIVQDTEBRLVDEVPVN
 TGVYKLRKLPDEFTEPYSIVFAENDEPLVNGIEVNASFASFEFLVTVNVEVPVP
 QOIFOANVSEDTAIGTKVNTARDPEGLVSYSLINDNKRGMKIDSVNCTSTAPL
 DRTESVYRQVAVTEVGSLSSTAYFHLVMDVNDNPRRLAKDYTGCFRCHPLSP
 GSILFEATDDQDSVRPKPTFALGRSLSDVSVSKINQTHRLSTHTRFEEQYD
 IPIINDGQGPNEGIVLSLVTFCCQCDGSCFPAGNQVGIPTVGMAVGLLTFTFLVI
 GIIIAVFIRMKQVEDPQSPENKPLRS"
 FEATURES
 source

polyA_signal 3605..3610
 BASE COUNT 1023 a 835 c 811 g 959 t
 ORIGIN
 Query Match 53.0%; Score 1774.4; DB 10; Length 3628;
 Best Local Similarity 80.4%; Pred. No. 0;
 Matches 2108; Conservative 0; Mismatches 501; Indels 13; Gaps 2;

9 TCTGACCACTGAATGGAAGAAAAGACTTTTAACCAACATTTTGTGACTTACAGAAG 68
 247 TCTCAGGGCTGAATGACGACAAACGACTTTTAAGAGTATTTTCTTACTTACGAAG 306
 69 AATTGAATTAAGAAAATCTATGATCTTACAGCCCATCTTCACTCCCTGTCTTCTAT 128
 307 AGTCCAGATGAAGAGC - AACATGATCTCTGCCCACCTTCACTTCTGTCTTCTAC 365
 129 GCTTATTTGGCACTGATATGAGCAAGAGGGAAGTTTATAGTGAACCCCTGAAGCCAT 188
 366 GCTTATTTGACGCGCATATGCGCAAGAGGGAAGTTTCAAGCGCCCTGAAGCCAT 425
 189 GACATTTCTATTTATGAGAGCCCAAGACGATCAATTAATTCCAGTTTAAGCCAA 248
 426 GACATTTCTATTTTGAAGCCCAAGACGATCAATTAATTCCAGTTTAAGCCAA 485
 249 TCTCTGCTGTGACTTTTGAATTAAGTGGAGACAGACAAATTTTGTATGATGAC 308
 486 CCTCCAGCTGTGACTTTTGAAGTGAACAGAGAGACGATGTATTTTAAGATGAAA 545
 309 GAGAGGACTTCTGATTAACAAGAGCTTGGACAGAGGAAACAAGATCTACATCT 368
 546 GGAATGACTTCTTATATCAAGAGATCTGGAACAGAGAAACAAGGCGGTTCAATCT 605
 369 CCAAGTTGACCCCTGACCTTAATGGAATTAATGAGAGGATCAATCTTACCAT 428
 606 GCAAGCTTGACCTTGAATTTCTAGAGAGCTATATGATGATGAGCCCACTCCCATCAT 665
 429 AGAATGAGAGCAATCAACAACATGAGCCACGTTTCTCAAGTCAAAAGTACGAAGCTC 488
 666 AGAATGAGAGCAATCAACAACATGAGCCACGTTTCTCAAGTCAAAAGTACGAAGCTC 725
 489 AGTAAGCAGAACTCTTCCGCAAGAAAGCCCTTCTTGTATGTCAATATCCACAGACTGA 548
 726 AGTGAAGCAGAACTCTTCCGCAAGAAAGCCCTTCTTGTATGTCAATATCCACAGACTGA 785
 549 TGATCCGCGCACTCCCAATGAGCCGTTTATTAACAGATTTGATCCAGCTCCCATGAT 608
 786 TGATCCGCGCACTCCCAATGAGCCGTTTATTAACAGATTTGATCCAGCTCCCATGAT 845
 609 CAACATATGATATCTTTCAGATCAACAAACAGGAGCCATCTCTTACCCGAGA 668
 846 CAACATATGATATCTTTCAGATCAACAAACAGGAGCCATCTCTTACCCGAGA 905
 669 GGGATCTCAGGAATTTGAATCTGTGAAGATCTCTTCTTAAATCTGTGATCTCAGTGA 728
 906 AGGATTTCAAGATTTGATCAATTAAGAAATCCGATCTACACCTGTGTCTCTGTGA 965
 729 GGAATGAGAGGCGCAGAGTGAATTTCTTCAAGTATACCAATCTGTGATCAATGAT 788
 966 GGAATGAGAGGCGCAGAGTGAATTTCTTCAAGTATACCAATCTGTGATATTAAGT 1025
 789 GACAGAGATATTTTGAAGACCAACAAACCTGTGAGATGTGAAAACCTCAACTGATCC 848
 1026 CAGAGAGAAATCTGGAAGACCAACAGAGCTGTGAGATTTGAGAAAACCTTAATGATCC 1085
 849 TCAACCCCATTAATCACTCAGGTGCGGTGAATGATCCGATGACATATATTCCTTAA 908
 1086 TCAACCCCATTAATCACTCAGGTGCGGTGAATGATCCGATGACATATATTCCTTAA 1145
 909 TGACAAAGAGAGCTGCAAGATTTCCATTTTCAATTTGACAGAGAGAGATTTTAACT 968
 1146 CAACAAAGAGAGCTGCAAGATTTTCAATTTTCAATTTGACAGAGAGAGATTTTAACT 1205
 969 GACTGAGCCCTTGACCGAGAGAAAAGATGATATGTTTTTATGCAAGTTGCAAAAG 1028

1206 GACTGAGCCCTTGTGACCGTAGAGAAAAGACTCTCATGTTTCTTGTCCACTCCAGAGA 1265
 1029 TGAGTACGGAAAACCACTTTCATATCCGCTGGAAAATTCATAGTAAAGTTAAATTA 1088
 1266 TGGAATGAGAAAACCTTGTGATATCCATGGAATTCGTGAAAAGTTATTAATTA 1325
 1089 TGATTAATCCACTACATATGTCCTGACAGTAAACCGTATTTGAGAGTCCAGAGATGAAC 1148
 1326 CGAACACCCACGACATGTCTCTCAAGTACGTATTTGAAGTCCAGAGATGAAGT 1385
 1149 ACTGGGTACAGTATGGGACCTTTAATGACATGACAGGATGAAGAAAATTAATGCCAA 1208
 1386 CTGGGTAGTATGAGATTTGGATCTTTGTGCTGCCACGATATGATGAAGTAAATTA 1445
 1209 CAGTTTCTTAATTAAGATTTGTGAGCAAACTCCCAACTCCCATGATGATGATCTT 1268
 1446 CAGTATTTTGAAGTATGACAGCTTTGTAGACCAACCAAGTTCCCTCAGATGAATCTT 1505
 1269 CTAATTCAAAACCTATGCTGGAATGTTACAGTTAGCTAAACAGTCTTGAAGAGCAGA 1328
 1506 TCTCATTTAGAGATATGAGAGAAAGTCCAGTTGGTAAACGTTCTTGAAGAGCAGA 1565
 1329 TACTCTCAGTACAACTTAACGATAGAGGTGTCTGACAAAGATTTCAAGACCTTTGTT 1388
 1566 CAGTCTCAGTACAACTTAACGATAGAGGTGTCTGACATGATTTCAAGACCTCTGTT 1625
 1389 TGTGCAATACAGTATATGATATCAATGATCAGATCCCATCTTTGAAAATCAATTA 1448
 1626 TCTTCAAGTCAAGTATATGATATCAAGATCCAGATTTCCATCTTTGAAAAGTCAAGCTA 1685
 1449 TGAAGAACCTGACCTTCTGTAAGACAAACATTTGGGTCCACCATCTTAACATCCAGG 1508
 1686 TGAAGAACCTGACCTTCTGTAAGACAAACATTTGGGTCCACCATCTTAACATCCAGG 1745
 1509 CACTGATGCTGATGAGCCATTAAGTGAAGTTCTAAATCTGTATCATATCAATTAAGG 1568
 1746 TACTGATGATGATGAGCCCTTTACCGGAGCTTAATTAATCTGTACAAAGTTGATACAGG 1805
 1569 AGACAGTGAAGGAGCGCTGGGGTTGACACATCCCATCAACACCGGATATGTCAT 1628
 1806 AGACAGTGAAGGAGCGCTGAAGTTGTCAAGATCCCATCAACATCAAGATATGTCAT 1865
 1629 AATTAAGAACCTCTTGAATTTGAACAGAGGTGTTCAACATTTGTGCAAGTCAAG 1688
 1866 GATCAAGAACCTCTTGAATTTGAACAGAGGTGTTCAACATTTGTGCAAGTCAAG 1925
 1689 AATCTGAGCCCTTGAATTTGTGTGTAAGTACAAATGCAAGTCTTTTGCCAAAGTTCA 1748
 1926 AATCTGAGCCCTTGAATTTGTGTGTAAGTACAAATGCAAGTCTTTTGCCAAAGTTCA 1985
 1749 GCTTATTTGACAGATGTAAGTGAAGACCTCAATTTTCCAAACAGTATCCAGCGAA 1808
 1986 GCTGACCGGTGACAGATGTAAGTGAAGTGTGATTTCCCAAGCAATATCCAAAGCGAA 2045
 1809 AGTCAAGTGAAGTATGATGTAAGTGAAGTGTGATTTCCCAAGCAATATCCAAAGCGAA 1868
 2046 GCTTACGAGAGATCTGCTATAGGCACTAAAGTGGCACTGTGACGCAAGGATCCGA 2105
 1869 AGTCTGACATTAAGTATTTCACTGAGGGAGACAAAGAGTTGGCTTTAAATTGACCA 1928
 2106 AGGCTGACATTAAGTATTTCACTGAGGGAGACAAAGAGTTGGCTTTAAATTGACCA 2165
 1929 CGTGACTGTGATGATTTTATGTTGCTCCATTTGACAGAGAACCGGAAGTCCATATG 1988
 2166 CGTGACTGTGATGATTTTATGTTGCTCCATTTGACAGAGAACCGGAAGTCCATATG 2225
 1989 GGTACAGGTGAGGCGACAGAGTGAAGGAGGCTTCCCTTAAGCTGTGATCAGAGTTCA 2048
 2226 AGTACAGGTGAGGCGACAGAGTGAAGGAGGCTTCCCTTAAGCTGTGATCAGAGTTCA 2285
 2049 CCTGATCTTATGATGATGATGAACACCTTCCAGGCTAGGCAAGATCAAGGGCTT 2108

Db	Accession	Source	Location/Qualifiers
Db	2286	CCTGTGCTCATTGATGGATGATGACAAACCCCTCGCTTACGTAAAGACTACACAGGCTT	2345
Qy	2109	GTTCTTCTTGCATCCCTCAGTGCACTGGAAGTCTCATTTTCAGAGCTTATGTATGA	2168
Db	2346	GTTCTTCTGCAATCCCTTAAGTGTGCCCCGAGAGCTCATTTTAAAGCACATGATGATGA	2405
Qy	2169	TCAGCACTTATTTGGGGGTGCCCATTTTACATTTTTCCTCGGACGTGGAACTTACAAA	2228
Db	2406	CCACAGTCAGTTCGAGAGGCCCAAGTTTACGTTTGCCCTTTGACAGAGAACTTACAAAG	2465
Qy	2229	CGACTGGGAATTTTCCAAATCAATATGATCATATGCCCCGACTGTCTACAGGACACAGA	2288
Db	2466	TGACTGGGAATGTGTCAAATATCAATATGACACATATGTATGACTCTTCACCAAGCACACG	2525
Qy	2289	CTTTGAGGAGAGGGCGTATGTCTGTTGATCCGCATCAATATGGGGGTCCGCCACCTT	2348
Db	2526	CTTTGAGGAGCAAGTTTATGATCATCTCCATCCTCATATGATATGGGGGAGCAACCAT	2585
Qy	2349	GGAAGGCAATGTCTTTTACCAAGTACATCTGCAGTTGTGTGAGAGGAAGATGTTCGG	2408
Db	2586	GGAAGGCAATGTCTCTCTTATCATAGTACTTTCGCAAGTGTGACGGAAGTGTTCG	2645
Qy	2409	GCCAGCAGGTACCCAGCATGGGATACCCACCTGTGGGACATGGCAGTGTATCTGTCAC	2468
Db	2646	GCCACAGGACCAACAGGTGGGATACCCACCTGTGGGACATGGCAGTGTATCTTCTGAC	2705
Qy	2469	CACCTTCTGTGTATGGTATTAATTTTACAGAGTTGTGTTATCCGCATTAAGAGATTA	2528
Db	2706	CACATTTCTGGTCATGTGATTAATTTTACGGTGTATTCATCCGATGA-----	2755
Qy	2529	AGGCAAAATATATGTTGAAAGTCTCAAGCATCTGAAGTCAAACTCTGAAGCTGAT	2588
Db	2756	--GGAAATATTAAGTTGAAGATCCACAGTCCCTGAGATTAAGCTCTGAGAGCTGAT	2813
Qy	2589	TTGAAAGAAATGTTGAATTATTAATTAAGCAAGTCTATTTC	2630
Db	2814	TTGAACAGAAATGTTCAAGGTTGTACACCAAGCACTGATCA	2855
RESULT	13		
Db	D87912	3414 bp	mRNA linear ROD 06-OCT-2000
LOCUS	D87912		
DEFINITION	Mus musculus mRNA for BILL-cadherin, complete cds.		
ACCESSION	D87912		
VERSION	D87912.1		
KEYWORDS	GI:9392287		
SOURCE	cadherin.		
ORGANISM	Mus musculus (strain: BALB/c) foetal liver Prob cell cell_line:389		
REFERENCE	CDNA to mRNA.		
AUTHORS	Mus musculus		
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclirognathi; Muridae; Murinae; Mus.		
JOURNAL	Ohnishi, K., Shimizu, T., Karasuyama, H. and Melchers, F.		
MEDLINE	The identification of a nonclassical cadherin expressed during B cell development and its interaction with surrogate light chain		
REFERENCE	J. Biol. Chem. 275 (40), 31134-31144 (2000)		
AUTHORS	2 (bases 1 to 3414)		
TITLE	Ohnishi, K.		
JOURNAL	Direct Submision		
MEDLINE	Submitted (18-SEP-1996) Kazuo Ohnishi, National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-Ku, Tokyo 162, Japan (E-mail:ohnishik@nih.go.jp)		
REFERENCE	Tel:03-5285-1111(ex.2133), Fax:03-5285-1150)		
AUTHORS	Location/Qualifiers		
TITLE	1. 3414		
JOURNAL	/organism="Mus musculus"		
MEDLINE	/strain="BALB/c"		
REFERENCE	/db_xref="taxon:10090"		
AUTHORS	/cell_line="389"		
TITLE	/cell_type="Prob cell"		
JOURNAL	/tissue_type="foetal liver"		
MEDLINE	93. .2576		
REFERENCE	CDS		

[illegible]

QY 728 AGGATGGAAGGCGAGAGTGAATTCCTTCACTGATACCACTCTGTGATATCTAG 787
 Db 730 AGGATGGAAGGCGAGAGTGAATTCCTTCACTGATACCACTCTGTGATATCTAG 789
 QY 788 TGACAGAGATATTTGGGAAAGACCAAACTCTGTGAGATGTGTGAAAACCTCACTGATC 847
 Db 790 TTGAGAGAGATATCTGGAAGACCAAACTCTGTGAGATGTGTGAAAACCTCACTGATC 849
 QY 848 CTCACCCCATCAAAATCACTCAGGTGCGGTGAATGATCCCGGTGCAAAATATCTCTAG 907
 Db 850 CTCACCCCATCAAAATCACTCAGGTGCGGTGAATGATCCCGGTGCAAAATATCTCTAG 909
 QY 908 TTGACAAAGAGAGAGCTGCAAGATTTCCATTTTCAATTTGACCAAGAGAGATATTTACG 967
 Db 910 TCAACAGAGAGAGCTGTGCGCGTTCCTCATCTGATGACCAAGAGAGATATTTATG 969
 QY 968 TGACTCAGCCCTTGGACCGGAGAGAAAAGATGATATGTTTTTATGCACTTGCAAGG 1027
 Db 970 TGACTCAGCCCTTGGACCGGAGAGAAAAGATGATATGTTTTTCTTTGCAACTGCGCAAGG 1029
 QY 1028 ATGATGAGGAGAAACCACTTTCATATCGGCTGGAATTCATGTAAAGTTAAAGATATTA 1087
 Db 1030 ATGATGAGTGAAGAAACCACTTTCATATCGAATTTATGTGAAGTTATGTACATTA 1089
 QY 1088 ATGATATCCACCTACATGTCCGTCAACAGTAAACCGTATTTGAGTCCAGAGATGAAAC 1147
 Db 1090 ATGACAAACCCAGCAATGTCTGTCTCCAGTACTGTATTTGAAGTCCAGAGATGAAAC 1149
 QY 1148 GACTGGGTATCAGTATGTGGACCTTACTGCACTGACAGGAGTGAAGAAATATCTGCCA 1207
 Db 1150 CATTTGGGTAAACAGATCGGATCTTTGAGGCCCATGTATGTAGCAAGCTATCAACATCA 1209
 QY 1208 ACAGTTTCTAAACCTACAGAGTTGTGAGCAACTCCCAACTCCCATGTAGATGATCT 1267
 Db 1210 ACAGTATTTGAAATACAGCTGTGAGCAACCAACCAAGTTCCTCAAGTGAATTT 1269
 QY 1268 TCTTAATCCAAACCTATGTGTAATGTATACAGTTAGCTTAAACAGTCTTTGAAGAGAG 1327
 Db 1270 TTCTCATTTGTTGAATATGAGGAGAAAGTTCAAGTTAAACAGTCTTTGAAGAGAGAG 1329
 QY 1328 ATATCTCTCAGTACAACTTAAAGATGAGGTGTCTGCAAAAGATTTCAAGACCTTTGTT 1387
 Db 1330 ACAGTCTCTCAGTACAACTTAAAGATGAGGTGTCTGCAAAAGATTTCAAGACCTTTGTT 1389
 QY 1388 TTGTGAAATCAAGCTTATGTATGATATCAATGATATCCCATCTTTGAAAATCAAT 1447
 Db 1390 ATATCTCAGTACAACTTATGTATGATATCAATGATATCCCATCTTTGAAAATCAAT 1449
 QY 1448 ATGGAACCTGACTCTTGTGTAAGACAAACATTTGGGTCCACATCTTTAACATCAAG 1507
 Db 1450 ATGGAAGCAAGACTCTGTCTGTAAGACAAACATTTGGGTCCACATCTTTAACATCAAG 1509
 QY 1508 CCACTGATGCTGATGAGCAATTTACTGGAGTTCTAAATTTCTGTATCATATCAATAAG 1567
 Db 1510 CTACAGATGCTGATGAGCCCTTTACAGGAGCTCTAAATCTGTACAAAGATTGTACAG 1569
 QY 1568 GAGACAGTGAAGGAGCGCTGGGGTTGACACAGATCCCATACCAACACCGGATATGTCA 1627
 Db 1570 GAGACAGGAGGAGGAGCTGAGAGTTGTCAAGATCCCAACCAATGACGAGATATGTCA 1629
 QY 1628 TAATTAAGAAAGCTCTGATTTTGAAGACAGCAGCTGTTTCCACATTTGTTCAAGAG 1687
 Db 1630 AGATCAAAAGAGCTCTTGTATTTGAAGACCAACAGTATGACATTTGTTTCCAGAG 1689
 QY 1688 AAAATCTGAGCTCTAGTGTGTTGGTGAAGTGAAGTGAAGTCTTTTTCAGAGTTCA 1747
 Db 1690 AAAATCTGAGCTCTAGTGTGTTGGTGAAGTGAAGTGAAGTCTTTTTCAGAGTTCA 1749
 QY 1748 CGCTTATTTGAGAGATGTAAGTGAAGCAGCTCAATTTTCCCAACAGTATTTCCAGAG 1807
 Db 1750 AGCTGATTTGAGAGATGTAAGTGAAGTGTGTTTCCCTCCAGCAAGATATTTCCAGAG 1809

QY 1808 AAGTCATGAGAGATGATCTATAGGACTAAAGTGGCAATGTGATCTGCCAAGATTCAG 1867
 Db 1810 AAGTCATGAGAGATGATCTATAGGACTAAAGTGGCAATGTGATCTGCCAAGATTCAG 1869
 QY 1868 AAGTCTGACATTAAGCTATTTCACTGAGGGGAGACACAGAGTTGGCTTTAAATTGACC 1927
 Db 1870 AAGGCTGACATGAGATTTATTTCACTGAAAGCAATATGAGAGCTGGCTTTAAATGACT 1929
 QY 1928 AAGTCACTGATGATCTTTAGTGTGCTTCATTTGACAGAGAGCGGAAGTCAATATC 1987
 Db 1930 CTGTCACTGATGATATTTATTTAGTGTGCTTCCTGATGAGAGAGAGAGAGTGTATC 1989
 QY 1988 GAGTCAAGATGTGTGAGCAAGAGTGAAGGAGGAGTCTTCTTAAGCTCTGTGTCAAGTTC 2047
 Db 1990 GAGTCAAGATGTGTGAGCAAGAGTGAAGGAGGAGTCTTCTTAAGCTCTGTGTCAAGTTC 2049
 QY 2048 AACTGATCTTATGATGTGATGACACACCTTCCAGGCTAGCCAAAGACTACAGGCT 2107
 Db 2050 AACTGATCTTATGATGTGATGACACACCTTCCAGGCTAGCCAAAGACTACAGGCT 2109
 QY 2108 TGTCTTCTGCGCATCCCTCAGTGCACCTGGAAGTCTCAATTTTCGAGGCTACTGATGAG 2167
 Db 2110 TGTCTTCTGCGCATCCCTCAGTGCACCTGGAAGTCTCAATTTTCGAGGCTACTGATGAG 2169
 QY 2168 ATCAGACATTAATTCGAGGCTCCCATTTTACATTTTCCCTCGGACAGTGAAGCTTACAA 2227
 Db 2170 ACCAGACATTAATTCGAGGCTCCCATTTTACATTTTCCCTCGGACAGTGAAGCTTACAA 2229
 QY 2228 ACAGCTGGGAAAGTTTCCAAATCAATGTATCTGATGCCGAGCTGTCTTACAGGACACAG 2287
 Db 2230 GTGACTGGGAAAGTTTCCAAATCAATGTATCTGATGCCGAGCTGTCTTACAGGACACAG 2289
 QY 2288 ACTTTGAGAGAGGAGGTATGTCTGTATCCGATCAATATGAGGGGTCCGACCT 2347
 Db 2290 GCTTTGAGAGAGGAGGTATGTCTGTATCCGATCAATATGAGGGGTCCGACCT 2349
 QY 2348 TGGAGAGCAATGTTTCTTTTACAGATTTACATTTCTGAGTGTGGAAGAAAGTTGTTCC 2407
 Db 2350 TGGAGAGCAATGTTTCTTTTACAGATTTACATTTCTGAGTGTGGAAGAAAGTTGTTCC 2409
 QY 2408 GGCAGCAGAGTACACAGACTGGGATACCACTGTGGGCAATGCAATGTGTATCTGTA 2467
 Db 2410 GGCAGCAGAGTACACAGACTGGGATACCACTGTGGGCAATGCAATGTGTATCTGTA 2469
 QY 2468 CCACTCTTGTGATGTGTATTTAGCACTGTGTATTTCCCATTAAGAGAGATA 2527
 Db 2470 CCACTCTTGTGATGTGTATTTAGCACTGTGTATTTCCCATTAAGAGAGATA 2529
 QY 2528 AAGGCAAGATATGTTGAAGTGTCAAGCATCTGAAGTCAAACTCTGAGAGCTGA 2587
 Db 2521 ---GGAAGATTAAGTTGAAGATTCGCACTCCCTGAGATTAAGCTCTGAGAGCTGA 2577
 QY 2588 TTTGAAAGAAATGTTGAATTTATATAGCAAGTGTATTTCA 2630
 Db 2578 TTTGAAAGAAATGTTGAATTTATATAGCAAGTGTATTTCA 2620

RESULT 14
 AF177669 3351 bp mRNA linear ROD 05-JUN-2001
 LOCUS AF177669
 DEFINITION Mus musculus liver-intestine cadherin mRNA, complete cds.
 ACCESSION AF177669
 VERSION AF177669.1 GI:5762489
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
 REFERENCES
 1 (bases 1 to 3351)
 Angres, B., Kim, L., Jung, R., Geesner, R. and Tauber, R.
 LI-cadherin gene expression during mouse intestinal development
 Dev. Dyn. 221 (2), 182-193 (2001)
 MEDLINE 21270055

PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

113/6485
2 (bases 1 to 3351)
Angres, B., Kim, L. and Tauber, R.
Direct Submersion
Submitted (14-AUG-1999) Institut fuer Klinische Chemie und
Biochemie, Virchow-Klinikum der Humboldt-Universitaet zu Berlin,
Augustenburger Platz 1, Berlin 13353, Germany
Location/Qualifiers

FEATURES

source

CDS

1. .3351
/organism="Mus musculus"
/strain="Swiss Webster/NIH"
/db_xref="taxon:10090"
/tissue_type="intestine"
/dev_stage="17 day embryo"
44. .2527
/note="LI-cadherin; cell-cell adhesion molecule"
/codon_start=1
/product="liver-intestine cadherin"
/protein_id="AA051125.1"
/db_xref="GI:5762490"

/translation="MWSAQLHFLCLTLVTCGYGEKFGSPKPMPTSSIFEGQEPS
QVIFQKTNPAVTFELGETDGI FKEKGLHYTRALDRFRVAHLQALADSHG
AIVDGPVITIEVDINDNRPLOSKEYGSRQNSRPGKPMYVATDIDPATNG
OLFOYVIOLOIINDWYFOIDSKTGAISLTPGSOELDKVMSYLVVSDMGQ
SENSFDITTYVDISIRENTKAPBVRIRNSDPRPHIKITTOVOMDPAOVSLVANK
KLSPPSPSIDQENITVQALDREKSHVFPAADENKPLAFLIETIKYIDIND
NPTCLSPVTFEVOENEPDGNISIGIFEAHMDENANNINILKYKVDQPKVPSGL
FLIGEGKVLQSKQKODSPQYNLSIVSDVDCTCYIQVADTEGLRIVDPTN
SNYKSLSEDTAIGSTIIIOATDEPFGSSKILYKIQVADTEGLRIVDPTN
AGYVKKIKPDLFPTQPVSTVFOANENEPVKGIEVYASFAPELLIVDNEVPEP
QRIFOANVEDDAVGRVAVTARDPEGLVYSISIKNMGCMKISVGEISAPL
DRETSVYRQVAVTEVSSLSSTADPFLHVLVDNDNPRKADYTGFCFPLSAP
GSLIFETVDDQOOLSRPKFTFALGREGLDWEVSKINDTHALSRHFRSEOVYN
IPRINDGQPPMEGTVEFLPVTFCQCEVSGCFRAGRODGIPTVGMVAGILTPFLVI
GILAVVFIRMKDKVENPOSPEKPLRS"

BASE COUNT
ORIGIN

966 a 748 c 746 g 891 t

Query Match 52.3%; Score 1747.8; DB 10; Length 3351;
Best Local Similarity 80.4%; Pred. No. 0;
Matches 2076; Conservative 0; Mismatches 492; Indels 13; Gaps 2;

QY 50 TTTGTGACTTACAGAAAGAAATTGGAATAAGAAACATGATCTTCAGGCCCATCTTC 109
DB 4 TTTCTTACTTACTGAAAGAGTCCAGTGAAGAGC-AAGATGCTGTCTGCCAGCTTC 62
QY 110 ACTCCGCTGCTCTTATGCTTATTTGGCACTGATATGCGCAAGAGGGAAGTTTA 169
DB 63 ACTTCTGTGCTCTTCTACACTTATTTGACCTGTGATATGGGAAAGGAAGTTTCA 122
QY 170 GTGAGCCCTGAAACCCATGACATTTTCTATTATGAAGCCCAAGAACCAATTAA 229
DB 123 GCGGTCCCTGAAAGCCCATGACATTTCTCAATTTTGAAGGCCAAGAACCGAGTCA 182
QY 230 TATTCAGATTAAAGCCCATCTCTGCTGTGACATTTTGAACCTAACTGGGAGACACA 289
DB 183 TATTCAGATTAAAGCCCAACCTCCAGCTGTGACATTTTGAAGCCGGAAGACAGATG 242
QY 290 ACATATTGTGATGAACGGAGGAGACTTCTGTATTACACAGAGCCCTTGACAGGAAA 349
DB 243 GTATATTTAAGATTGAAGAAAGATGACTTCTGTATCACAAAGAGCCCTGACAGAGAAA 302
QY 350 CAAGATCTACTACAAATCTCCAGGTGACAGCCCTGAGCGCTTAATGAATTAATGAGAG 409
DB 303 CAAGAGCGGTTCATCATCTGACAGCTTGACGCTTGATTTCTCATGAGACTAATGAGAGC 362
QY 410 GTCCAGTCCCTATCACTAAGTGAAGAGACATTAAGACATGACATCCAGCTTTCTCC 469
DB 363 GTCCAGTCCCTATCACTAAGTGAAGAGACATTAAGACATGACATCCAGCTTTCTCC 422
QY 470 AGTCAAAAGTGAAGAGGCTCAGTGAAGGAGAACTCTCGCCCAAGAAAGCCCTTCTGTATG 529
DB 423 AGTCAAAATATGAAGGCTCAGTGAAGGAGAACTCTCGCCCAAGAAAGCCCTTCTGTATG 482

QY 530 TCAATGCCACAGACCTGAGTATCCGGCCACTCCCAATGGCCAGCTTTATACAGATG 589
DB 483 TCAATGCCACAGATCTGAGTATCCGGCCACTCCCAATGGCCAGCTTTATACAAATG 542
QY 590 TCATCCAGCTTCCATGATCAACATGTCATGATCTTTCAGATCAACAAACGGGAG 649
DB 543 TCATCCAGCTTCCCAATCAACGATGTATGATCTTTCATATGACAGCAAAACGGGAG 602
QY 650 CCATCTCTTTACCCGAGAGGATCTCAGAAATTGAATCTCTCTAAGATCTCTCTATA 709
DB 603 CATATCGCTTACCCAGAGAGATCCAGAAATGGATCCAGTGAAGAAATCTCTCTATA 662
QY 710 ATCTGTGATCTCAGTGAAGACATGGAGGCCAGATGAGATCTCTCAGTATACCA 769
DB 663 ACCTGTGTGCTCTGAGTGAAGACATGGGGGCCAGATGAGATCTCTCAGTATACCA 722
QY 770 CATCTGTGATATCATAGTGAAGAAATATTTGAAAGCCACCAAACTGTGAGATG 829
DB 723 CATATGTATATTTCTATTTAGAGAAATATCTGAAAGCACACAGAACCCGTGAGATTA 782
QY 830 TGGAAATCTCAACTGATCTCTACCCCATCAAAATCATCAGGTGCGGTGAGATGATCCG 889
DB 783 GAGAAATCTCGACTGATCTCTACCCCATCAAAATCATCAGGTGAGTGAATGACCCAG 842
QY 890 GTGCAAAATATCTCTAGTTGAACAAAGAGGCTGCAAGATTTCCATTTTCAATTGACC 949
DB 843 GGGCCAGATTTCTCTAGTTGAACAAAGAGAGAGTGTGCGGCTTCCATTTCTGATGACC 902
QY 950 AGAAGAGAAATATTTAGTACTCAGCCCTTGACCCGAGAGAAAGATGATATGTT 1009
DB 903 AGAAGAGAAATATTTAGTACTCAGCCCTTGACCCGAGAGAAAGATGATATGTT 962
QY 1010 TTTATCAGTTCAGAAAGATGAGTACGAAACCACTTTTCAATCTCGCTGGAATTCATG 1069
DB 963 TCTTTCAGACTGCCAAGATGAGAAATGAGAAACCACTTTCATATTCAGTGAATTTATG 1022
QY 1070 TAAAGTTAAAGTATTATGATTAATCAACCTTACATGCTCGTACAGTAAACGATATTTG 1129
DB 1023 TGAAGTTATGATTAATTAATGACACCCAGACATGCTGTCTCCAGTACGATATTTG 1082
QY 1083 AAGTCCAGAGATGAGAACCATTTGGTAAACGATTCGGGATCTTTGAGGCCCATGATATG 1142
DB 1190 ATGAGAAATATCTGCCAAGCTTTTCTAACTACAGATTTGGAGCAAACTCCCAAC 1249
QY 1143 ACGAAGCTAACCAATCAACATGATTTTGAATTAACAAGCTTGTAGACCAAAACCCAAAG 1202
DB 1250 TTCCCATGATGAGACTCTTCTAATCCAAACCTATGCTGAGATGTTACAGTTAGCTTAAC 1309
QY 1203 TTCCCTCAGTGAAGCTTTTCTCATTTGGTGAATATGAGGAAAGTTCAAGTTAAGTAAAC 1262
QY 1310 AGTCTTGAAGAGCAAGATACTCTCAGTACAACTTAACATGAGAGGTGTGACAAAG 1369
DB 1263 AGTCTTGAAGAGCAAGAGAGTCTCAGTACAACTTAACATGAGAGGTGTGACAAAG 1322
QY 1370 ATTTCAAGACCTTTGTTTGTGCAATCAAGTTATGATATCAATGATCAATCCCA 1429
DB 1323 ATTTCAAGACCTCTGTTATATTAAGTCAACGTTATGATATCAATGATCAATCCCA 1382
QY 1430 TCTTTGAAAATGAGATTAAGAAACCTGACTGTGCTGAGACACAAACATTGGGTCGA 1489
DB 1383 TCTTTGAAAATGAGATTAAGAAACCTGACTGTGCTGAGACACAAACATTGGGTCGA 1442
QY 1490 CCATCTTAACATCCAGGCACTGATGCTGATGAGCAATTAATGAGAGTTCTAAATTC 1549
DB 1443 CCATCTTAACATCCAGGCACTGATGCTGATGAGCAATTAATGAGAGTTCTAAATTC 1502
QY 1550 TGTATCATATCATTAAGAGAGACAGTGAAGAGCGCTCGGGGTTGACACAGATCCCATTA 1609
DB 1503 TGTACAGATTGTATCAGGGAGACAGCGAGGAGAGACTAGAAAGTTGTACAGATCCCA 1562

Db 301 CAGCTTTTATCAAAATGTCATCCAGCTTCCAAATCAACAATGTATGACTTCAG 360
 QY 631 ATCAACAACAAACGGAGCCATCTCTTACCCGAGAGGATTCAGAAATGATCT 690
 Db 361 ATGCAACAACAAACAGGGGCAATCTCCCTTACCCCAAGGTTCTCAGATATGATCA 420
 QY 691 GCTAAGAAATCCTTCTTAATCTGATCTCAGTGAAGCAATGGAGGCGAGAGAG 750
 Db 421 ATAAAGATCCTTCTCAACCTGGTGTCTCTGTGAAGCATGGGGGCGCAATGAG 480
 QY 751 AATTCCTTCAAGTATCAACATCTGTGATATCATAGTGAAGAAATATTTGAAAGCA 810
 Db 481 AATTCCTTCAAGTATCAACATCTGTGATATCATAGTGAAGAAATATTTGAAAGCA 540
 QY 811 CCAAAACCTGTGAGATGTGGAATACTCAACTGATCTTCAACCCATCAAAATATCTAG 870
 Db 541 CCAAGGCTGTGAGATGTGGAATACTCAACTGATCTTCAACCCATCAAAATATCTAG 600
 QY 871 GTGCGGTGAATGATCCCGGTGCAAAATATCTTATGTGCAAAAGAAAGCTGCCAAGA 930
 Db 601 GTGCAAGTGAATGACCAAGTGGCCCACTATCTTAACTCAACAGAGAAAGCTGCCAAG 660
 QY 931 TTCCCATTTTCAATTTGACCAAGAAAGATATTTACGTGACTCAAGCCCTTGAACCAAGA 990
 Db 661 TTTCATTTCTGATCGAACCAAGAAAGAAATTTACGTGACTCAAGCCCTTGAACCGTAG 720
 QY 991 GAAAAGATGATATGTTTTTTTATGCAAGTGGCAAGAGTACGAAACCACTTCA 1050
 Db 721 GAAAAGATGATCTCATGTTTTCTTTGCACTGCAAGCAAGCAAGAAATGAAAACCCCTTGA 780
 QY 1051 TATCCGCTGGAATTCATGTTAAAGTTAAAGATATTAATGATATTCACACCTACATGTCG 1110
 Db 781 TATCCACTGGAATTCGTTGAAATGTTATGACATTAACACACCAACGACATGTCTC 840
 QY 1111 TCACCACTAAACCGTATTTAGGTCCAGAGAAATGAACGACTGGTAAACATATCGAGACC 1170
 Db 841 TCTCAAGTGAATCTGATTTGAAGTCCAGAGAAATGAAGCTTGGTAGATGGGATC 900
 QY 1171 CTTACTGACATGACAGGAGTGAAGAAATATCTGCCAAGTTTTCTTAACTACAGAT 1230
 Db 901 TTTGCTGCCCAAGATGAGTGAAGCAACATATCAACATATTTGATGTACAGACT 960
 QY 1231 GTGGAGAACTCCCAAACTTCCATGATGATGACTCTTCCATCAAACTATGCTGA 1290
 Db 961 GTAGACCAACACCCAAAGTTCCCTCAGATGATCTTTCTCATGAGCAATATGAGAGA 1020
 QY 1291 ATGTTACAGTTAGCTAAACAGTCTCTTGAAGCAAGATATCTCTCAGTACCACTTAACG 1350
 Db 1021 AAGGTCCAGTTGGTAAACGGTCTTGAAGCAAGCAGTCTCAGTACCACTTAACA 1080
 QY 1351 ATGAGGTGTCTGCAAAAGATTTCAAGACCTTTGTTGTGCAAAATCAAGTTATGAT 1410
 Db 1081 GTAGAGGTGTCTGCAATGATTTCAAGACTCTGTGTCTTCAAGTCAAGTTATGAT 1140
 QY 1411 ATCATGATGATGATCCCATCTTTGAAAAATCAGATTAAGAACTGACTGCTGA 1470
 Db 1141 ATCAAGATCAGATTTCCATCTTTGAAAGATCAGACTATGAGCAAGACCTTTCTGA 1200
 QY 1471 GACACAAACATTTGGATCCACCATCTTAACATCCAGGCCACTGATGATGAGCAATT 1530
 Db 1201 GACACAGCATTCGATCCACCATCTTAATCATCCAGCTACTGATGATGAGCCCTT 1260
 QY 1531 ACTGGAGTTCTTAAATTTCTGTATCATATAAGGAGACAGTGAAGGACGCTGGGG 1590
 Db 1261 ACCGGAGCTTAAATTTCTGTATCAAGTTGTACAGGAGACACTGAGGGAGAGCTGAA 1320
 QY 1591 GTTACACAGATCCCAACACACCGGATATGTCATTAATAAAGCCCTTGAATTT 1650
 Db 1321 GTTGTCAAGATCCCAACACACCGGATATGTCATTAATAAAGCCCTTGAATTT 1380
 QY 1651 GAAAACAGAGCTGTTCCACATTTGTGTTCAAGCAGAAAAATCCTGAGCCTTATGTT 1710
 Db 1381 GAAAACAGAGCTTATCTAGATTTGTGTTCAAGCAGAAAAATCCTGAGCCTGTTGAT 1440

QY 1711 GGTGAGATGACAAATGCAAGTTCTTTGGCAAGTTCAAGCTTATGTGACAGATGTAAT 1770
 Db 1441 GGCATGAGATGACAAATGCAAGTTCTTTGGCTCTTTGAGTCAAGCTGAGATGTAAT 1500
 QY 1771 GAAGCACTCAATTTTCCCAACAGTATTCAGAGCAAGATGAGTGAAGTGAAT 1830
 Db 1501 GAGTGCCTGATTTTCCCAAGAAATATTCAGAGCAAGATGAGTGAAGTGAAT 1560
 QY 1831 GGCATTAAGTGGGCAATGATGATGCAAGATTCAGAAAGTGTGACATTAAGCTTCA 1890
 Db 1561 GGCATTAAGTGGGCAATGATGATGCAAGATTCAGAAAGTGTGACATTAAGCTTCA 1620
 QY 1891 CTGAGGGAGACACAGAGTTGGCTTAAATTTGACAGTGAAGTGAAGTGAAT 1950
 Db 1621 CTGAAAGACATTAAGAGAGTTGGCTTAAATTTGACAGTGAAGTGAAT 1680
 QY 1951 GTGCTTCATTTGACAGAGAGCCGGAAGTCCATATCGGATCAAGTGTGGC 2003
 Db 1681 ACTGCTCCGCTGACAGAGGAAACAGAAAGTGTGATCAAGTCAAGTGTGGC 1733

Search completed: June 21, 2003, 23:43:29
 Job time : 8339 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2003, 01:07:55 ; Search time 4860 Seconds
(without alignments)
4982.209 Million cell updates/sec

Title: US-10-025-380-1081
Perfect score: 4321
Sequence: 1 MIIQAHLSHSLIMLYATG.....DKGKNVESQAASEVKEPLRS 832

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Delop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBY=xlp
-Q=/cgm21/USFTO.spool/US10025380/runat_17062003_173328_24668/app_query.fasta_1.1031
-DB=GenBml -QPM=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=bloum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pico -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US10025380 @CGN 1.1 3745 @runat_17062003_173328_24668 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUTRY -NBS -SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenBml.*
1: gb_ba.*
2: gb_hng.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ro.*
12: gb_ey.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vl.*
30: em_htg_hum.*
31: em_htg_in.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rtd.*
36: em_htg_man.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4321	100.0	2459	6 I40307	I40307 Sequence 2
2	4321	100.0	2459	6 I81197	I81197 Sequence 2
3	4321	100.0	3345	6 AX193509	AX193509 Sequence
4	4321	100.0	3345	6 AX330272	AX330272 Sequence
5	4321	100.0	3345	6 AX332252	AX332252 Sequence
6	4321	100.0	3345	6 AX410642	AX410642 Sequence
7	4321	100.0	3345	9 HSU07969	U07969 Human Intes
8	4303	99.6	3654	6 AX330312	AX330312 Sequence
9	4303	99.6	3654	6 AX335495	AX335495 Sequence
10	4303	99.6	3654	6 AX411163	AX411163 Sequence
11	4303	99.6	3654	6 HSRNALICA	X83228 H. sapiens m
12	3450	79.8	3351	10 AF177669	AF177669 Mus muscu
13	3450	79.8	3351	10 D87912	D87912 Mus muscu
14	3414	79.0	3628	10 RNCADHER	X78997 R. norvegicu
15	2328	53.9	1733	10 RAPPDRPT	L46874 Rattus norv
16	1427.5	33.0	3373	5 AF428098	AF428098 Datio rer
17	954	22.1	2846	4 OC028945	U28945 Oryctolagus
18	922.5	21.3	2764	9 AF016272	AF016272 Homo sapi
19	922.5	21.3	2820	9 BC027912	BC027912 Homo sapi
20	908.5	21.0	2848	6 AX092366	AX092366 Sequence
21	908.5	21.0	2848	6 AX358934	AX358934 Sequence
22	908.5	21.0	2848	6 AX362427	AX362427 Sequence
23	878.5	20.3	2864	10 BC015251	BC015251 Mus muscu
24	878.5	20.3	2927	10 AF016271	AF016271 Mus muscu
25	878.5	20.3	2938	6 AX338110	AX338110 Sequence
26	877.5	20.3	2493	6 AX338112	AX338112 Sequence
27	862.5	20.0	2799	9 AK093905	AK093905 Homo sapi
28	758.5	17.6	2516	6 AX396194	AX396194 Sequence
29	715	16.5	2954	5 CHKRCG	D14459 Chicken mRN
30	709.5	16.4	2827	5 AF418565	AF418565 Datio rer
31	705	16.3	3230	10 MMRRCADA	X69966 M. musculus
32	705	16.3	3274	10 HSNRC	D14888 Mouse mRN
33	699.5	16.2	2853	9 HSNCAD	X54315 Human mRN
34	699.5	16.2	4132	9 S42303	S42303 N-cadherin
35	699	16.2	3690	5 S82457	S82457 Xenopus mat
36	696.5	16.1	3203	5 GGNCAD	X07277 Chicken mRN
37	694.5	16.1	3269	10 BC022107	BC022107 Mus muscu
38	694.5	16.1	4321	6 AX305854	AX305854 Sequence
39	694.5	16.1	4321	10 AB008811	AB008811 Mus muscu
40	694.5	16.1	4321	10 MUSCADNA	M31131 Mouse neuta
41	691.5	16.0	3448	9 HUMNCAOH	M34064 Human N-cad
42	688.5	15.9	3839	4 BTCDH	X53615 B. taurus mR
43	687.5	15.9	3310	9 AK091496	AK091496 Homo sapi
44	686.5	15.9	3036	10 AF097593	AF097593 Rattus no
45	686.5	15.9	4350	10 AB017695	AB017695 Rattus no

RESULT 1

ALIGNMENTS

140307
LOCUS 140307 2499 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 2 from patent US 5620855.
ACCESSION 140307
VERSION 140307.1 GI:2082599
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 2499)
AUTHORS Danczig, A.H., Hoskins, J.A. and Skatrud, P.L.
TITLE Mammalian influx peptide transporter
JOURNAL Patent: US 5620855-A 2 15-Apr-1997;
FEATURES
Location/Qualifiers
1..2499
/organism="unknown"
BASE COUNT 728 a 559 c 570 g 642 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2499
Score: 4321.00 Matches: 832
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-025-380-1081 (1-832) x 140307 (1-2499)
QY 1 MetIleuGlnAlaHisIleuHisSerLeuCysLeuLeuMetLeuTYrLeuAlaThrGly 20
DB 1 ATGTAATCTTCAGGCCCATCTTCACTCCCTGTCTCTTATGCTTATTTGGCAACTGGA 60
QY 21 TYrGlyGlnGlyGlyIlePheSerGlyProLeuIleProMetThrPheSerIleTYrGlu 40
DB 61 TATGGCCAGAGGGGAAAGTTAGTGAACCCCTGAACCCATGACATTTTCTATTATGAA 120
QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheIleValAsnProProAlaValThrPhe 60
DB 121 GGCCAGAGAACCGAGTCAATATATATTCAGTTTAAGGCCAATCCCTGCTGTACTTT 180
QY 61 GluLeuThrGlyGlyIleThrAspAsnIlePheValIleGluArgGlyGlyLeuTYrTYr 80
DB 181 GAATTAATGGGGAGACAGACACATATTTGTGATAGAACGGAGGACTTCTGTATTAC 240
QY 81 AsnArgAlaLeuAspArgIleThrArgSerThrHisAsnLeuGlnValAlaIleLeuAsp 100
DB 241 AACAGAGCTTGACAGAGGAAACAAGATCTACTCAATCTCCAGGTGACAGCCCTGGAC 300
QY 101 AlaAsnGlyIleIleValGlyGlyProValProIleThrIleGluValIleAsnSerArg 120
DB 301 GCTATATGAATTAATAGTGGAGGGTCCAGTCCCTATCACCATAGAAAGTAAGACATCAAC 360
QY 121 AspAsnArgProThrPheLeuGlnSerIleTYrGlyGlySerValArgGlnAsnSerArg 140
DB 361 GACAAATGACCCACGTTTCTCCAGTCAAGTACAGAGCTCAGTAAGGCAAACTCTCGC 420
QY 141 ProGlyIleProPheLeuTYrValAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
DB 421 CCAGAGAAAGCCCTTCTGTATGTCAATGCCACAGCTGATGATCCGGCAGCTCCCAAT 480
QY 161 GlyGlnLeuTYrTYrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTYrPhe 180
DB 481 GGCCAGGCTTTATACCAATGTCATCCAGCTTCCCATGATCAACAATGATGATGACTT 540
QY 181 GlnIleAsnAsnIleThrGlyAlaIleSerLeuThrArgGlyGlySerGlnGluLeuAsn 200
DB 541 CAGATCAACAACAAACGGAGCCATCTCTCTTACCCAGAGGAGATCTCAGAAATTGAAT 600
QY 201 ProAlaIleAsnProSerTYrAsnLeuValIleSerValIleAspMetGlyGlyGlnSer 220
DB 601 CTGTGTAAGATTCCTCTATATATCTGTGATCTCTGATGAGACATGGAGGCGACAGT 660

QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTYrIle 240
DB 661 GAGAAATCTCTTCAAGTATACCAATCTGTGATATATAGTGAACAGAAATATTGGAAA 720
QY 241 AlaProIleProValGluMetValGluAsnSerThrAspProHisProIleIleValThr 260
DB 721 GCACCAAAAACCTGTGAGATGGTGGAAAACCAACGATCTCCACCCCAACAAACACT 780
QY 261 GlnValArgThrAsnAspProGlyAlaGlnTYrSerLeuValAspIleGlyIlePro 280
DB 781 CAGGTCCGGGAGATATCCCGGTGCAATATCTTACTTGACAAAGAGAGCTGCCA 840
QY 281 ArgPheProPheSerIleAspGlnGlyIleAspIleTYrValThrGlnProLeuAspArg 300
DB 841 AGATTCCATTTTCAATTGACCGAGAGAGATATTAGTGACTCAGCCCTTGACCGCA 900
QY 301 GlyGluIleAspAlaTYrValPheTYrAlaValAlaIleAspGluTYrGlyIleProLeu 320
DB 901 GAAGAAAAGATGACATATGTTTTTATGCAAGTTCAGAAAGATGATCGGAAAACACTT 960
QY 321 SerTYrProLeuGluIleHisValIleValIleAspIleAsnAspAspProThrCys 340
DB 961 TCATATCCCGTGGAAATTCATGTAAAGTTAAAGATTAATGATATCACTTACATATG 1020
QY 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
DB 1021 CCCTCACCAAGTACCCATTTTATGAGTCCAGAGAAATGAACGACTGGTAACTATCCGG 1080
QY 361 ThrLeuThrAlaHisAspArgAspGlyGluAsnThrAlaAsnSerPheLeuAsnTYrArg 380
DB 1081 ACCCTTACGACACATACAGAGGATGAAGAAATACGCCAACAGTTTCTTAACTACAG 1140
QY 381 IleValGlnGluThrProIleIleProIleProMetAspGlyLeuPheLeuIleGlnTYrAla 400
DB 1141 ATGTGGAGAAATCTCCAAACTTCCATGAGATGAGACTTCTTATCCAAACTTATGCT 1200
QY 401 GlyMetLeuGlnLeuAlaIleAspIleSerLeuIleValGlyAsnProGlnTYrAsnLeu 420
DB 1201 GGAATGTACAGTACTTAAACAGCTCTTGAAGAACAGATCTCTGATCAACTTA 1260
QY 421 ThrIleGluValSerAspIleAspPheIleValThrLeuCysPheValGlnIleAsnValIle 440
DB 1261 ACGATAGAGGTGTCGACAAAGATTCAAGACCTTGTGTGCAAAATCAACGTTATT 1320
QY 441 AspIleAsnAspGlnIleProIlePheGluIleSerAspTYrGlyAsnLeuThrLeuAla 460
DB 1321 GATATCAATGATAGATCCCATCTTGAAGAAATCAGATTATGGAACCTGACTTGTCT 1380
QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
DB 1381 GAAGACACAAACATTTGGGTCCACATCTTAACCATCAGGCCCATGATGCTATGAGCCA 1440
QY 481 PheThrGlySerSerIleIleValIleValIleIleValGlyAspSerGlyIleValGlu 500
DB 1441 TTTACTGGAGATTCTAAATATCTGTATCATATATTAAGGAGAACAGTGAAGGACCCCTG 1500
QY 501 GlyValAspThrAspProHisThrAsnThrGlyTYrValIleIleIleIleValIleAsp 520
DB 1501 GGAGGTGACACAAATCCCATCCATCCACACCGGATATGCTAATTAAGAGCCCTTGAT 1560
QY 521 PheGluThrAlaAlaValSerAsnIleValPheIleValGluAsnProGluProLeuVal 540
DB 1561 TTTGAAACAGACACTGTTTCCACATTTGCTTCAAAGAGAAATCTGAGCCTCTAGTG 1620
QY 541 PheGlyValIleTYrAsnAlaSerSerPheAlaIlePheThrLeuIleValIleThrAspVal 560
DB 1621 TTTGGTGTGAAGTACATGCAATGCTTTTGGCAAGTTCAAGCTTATTTGACAGAGTGTG 1680
QY 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaIleValIleSerGluAspValAla 580
DB 1681 AATGAAGCACCTTAATTTTCCCAACAGTATTCAGGAAATCAGTGAAGATGTAGCT 1740
QY 581 IleGlyThrIleValIleGluAsnValThrAlaIleAspProGluIleLeuAspIleSerTYr 600

```

Db      1741 ATAGGCACTAAAGGGCAAGTCACTGCCAAGATCCAGAGGCTCGGACATTAAGCTAT 1800
Qy      601 SerleuArgGlyAspThrArgGlyTTPleuLysIleAspHisValThrGlyGluIlePhe 620
Db      1801 TCACCTGAGGGAGACACAAAGAGGTGGCTTAATAATTAACAACGAGCTGGGAGATCTTT 1860
Qy      621 SerValAlaProleuAspArgGluAlaGlySerProTyrArgValGlnValAlaIleThr 640
Db      1861 AGGTGGCTCATTTGACACAGAAAGCCGGAAGTCATATCGGTACAAAGTGGCCACA 1920
Qy      641 GluValGlyGlySerSerLeuSerSerValSerGlnPheHisLeuIleLeuMetAspVal 660
Db      1921 GAAGTAGGGGGGTCTTCTTAAGCTCTGTGTCAAGATTCACACGATCTTATGATGATG 1980
Qy      661 AsnAspAsnProProAspGluAlaValAspTyrThrGlyLeuPhePheCysHisProleu 680
Db      1981 AATGACAACTCCCAAGGCTAGCCCAAGACTACACGGGGCTTGTCTTGTCCATCCCTC 2040
Qy      681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheArgGly 700
Db      2041 AGTGCACTGGAAGTCCATTTTCGAGGCTACTGATGATGATGATGATGATTTCCGGGT 2100
Qy      701 ProHisPheThrPheSerSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerLys 720
Db      2101 CCCCATTTTACATTTTCCCTCGGCACTGGAAGCTTACAAAACGACTGGGAAGTTCCAAA 2160
Qy      721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluValAlaTyr 740
Db      2161 ATCAATGGTACTACTAGCCCACTGTCTTACCAAGGACACAACTTTGAGAGAGGGCTAT 2220
Qy      741 ValValLeuIleArgIleAsnAspGlyGlyArgProProleuGluGlyIleValSerLeu 760
Db      2221 GTGGTCTTGATCCGCATCAATGATGAGGGGTCGCGCACCTTGGAAGGCATTTGTTCTTA 2280
Qy      761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
Db      2281 CCAGTTACATTCGTGCAGATTGTGTGGAAAGAAAGTTGTTCCGGCCAGAGGTTCACAGACT 2340
Qy      781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
Db      2341 GGAATACCCATGCGGCATGAGGCGATGTTGATGATGATGATGATGATGATGATGATGAT 2400
Qy      801 IleIleLeuAlaValAlaPheIleArgIleIleLysAspLysGlyLysAspAsnValGlu 820
Db      2401 ATTAATTTTACAGTTGTGTTTATCCGATTAAGAGATTAAGGCAAAAGATATATTTGAA 2460
Qy      821 SerAlaGlnAlaSerGluValLysProleuArgSer 832
Db      2461 AGTGCTCAAGCATCTGAAAGTCAAACTCTGAGAAAGC 2496

RESULT 2
LOCUS      181197 181197 2499 bp DNA linear PAT 10-JUN-1998
DEFINITION Sequence 2 from patent US 5710018.
ACCESSION 181197
VERSION     181197.1 GI:3209487
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 2499)
AUTHORS     Danczig, A.H., Hoskins, J. Ann. and Skatrud, P.L.
TITLE       Mammalian influx peptide transporter
JOURNAL     Patent: US 5710018-A 2 20-JAN-1998;
            Location/Qualifiers
            source 1. 2499
            /organism="unknown"

BASE COUNT  728 a 559 c 570 g 642 t
ORIGIN
Alignment Scores: 0 Length: 2499
Pred. No.:

```

```

Score:      4321.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB:         6
US-10-025-380-1081 (1-832) x 181197 (1-2499)

Qy      1 MetIleLeuGlnAlaHisLeuHisSerLeuCysLeuLeuMetLeuTyrIleuAlaThrGly 20
Db      1 ATGATTACTTACGGCCATCTTCACTCCCTGTGTCTTATGATGATGATGATGATGATGATGAT 60
Qy      21 TyrGlyGlnGluGlyLysPheSerGlyProleuLysProMetThrPheSerIleTyrGlu 40
Db      61 TATGGCCAAAGAGGGGAAGTTTACTTGAGACCCCTGAACCCATGACATTTTCTATTATTAAGA 120
Qy      41 GlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAsnProProAlaValThrPhe 60
Db      121 GGCCAAAGAACCGAGTCAAAATTAATTCAGTTTAAGGCCAATCCTCCTGCTGACTTTT 180
Qy      61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluArgGlyGlyLeuLeuTyrTyr 80
Db      181 GAACATACTGGGAGACAGACAAATTAATTTGTGATGAACGGAGGACTTGTGATTAAC 240
Qy      81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
Db      241 AACAGAGCCTTGACAGGGAAACAAAGATCTACATCTCAGGTTGACGCCCTGGAC 300
Qy      101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValLysAspIleAsn 120
Db      301 GCTAATAGGATTAATAGGAGGGGTCCAGTCCCTATACCACTAAGAGTGAAGACATCAAC 360
Qy      121 AspAsnArgProThrPheLeuGlnIleSerLysTyrGlyGlySerValArgGlnAsnSerArg 140
Db      361 GACAAATGACCCAGTTTCTCCAGTCAAGTAAAGAGGCTCAGTAAGGACGAATCTCGC 420
Qy      141 ProGlyLysProPheLeuTyrValAlaAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
Db      421 CCAGGAAGGCCCTTCTTGTATGTCATGACACAGACTGGATGATCCGGCCATCTCCAAAT 480
Qy      161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
Db      481 GGCCACCTTATTAACAGATTCATCCAGCTCCCATGATCAACATGATCATGTAATCTT 540
Qy      181 GlnIleAspAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
Db      541 CAGATCAACAAACAAACCGGAGGCATCTCTTAACCCGAGAGGATCTCAGGAATTTGAAT 600
Qy      201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlyGlnSer 220
Db      601 CTGTCTAAGATCTCTTCTTAATCTGTGATCTCAGTGAAGACATGGAGGCCAGAGT 660
Qy      221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGlnAsnIleTyrLys 240
Db      661 GAGAAATTCCTTCAGTATACCAATCTGTGATATCATATGATGACAAAGAAATATTGGAAA 720
Qy      241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
Db      721 GCAACCAAAACCTGTGAGATGTGGAAACCTCAACGATCTCAACCCATCAAAATCACT 780
Qy      261 GlnValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGluLysLeuPro 280
Db      781 CAGTCCGGGTGAATATATCCCGGTGACAAATATCTTACTTGAACAAAGAGAGCTGGCCA 840
Qy      281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrValThrGlnProleuAspArg 300
Db      841 AGATTCCCAATTTTCAATGACCAAGAAAGATATTATTAAGTCAAGCCCTTGAGACCA 900
Qy      301 GluGluLysAspAlaTyrValPheTyrAlaValAlaValAspGluTyrGlyLysProleu 320
Db      901 GAAGAAAGAGATGATATGTTTTTATGAGATTGCAAAAGAGATGATGAGAAACCACTT 960
Qy      321 SerTyrProleuGluIleHisValLysValLysAspIleAsnAspAsnProProThrCys 340

```

Db	961	TCATATCCGCTGGAAATTCATGTAAATCTTAAAGATATTATATGATATATCCACTCAACAT	1020
OY	341	ProSerProValThrValPheGluValGlnGlnubengluArgLeuGlyAsnSerIleGly	360
Db	1021	CCGTACACAGTAACCGGTATTGAGGTCCAGGAGAAATGAACACACTGGGTAAACAGTATCCGG	1080
OY	361	ThrLeuThrAlaHisAspArgAsnGluGlnubenthAlaasnSerPheLeuAsnThrArg	380
Db	1081	ACCTTACTGTGACATGACAGGAGATGAAGAAATATCTGCACACAGTTTCTTAACTCAACGG	1140
OY	381	IleValGlnGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyzAla	400
Db	1141	ATTGGAGGCAAACTCCCAACTTCCATGATGACCTTCTCTATATCCAAACGTATCT	1200
OY	401	GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyzAsnLeu	420
Db	1201	GGAAATGTACAGTTAGTCTAAACAGCTCTTGAAGAGCAAGATACTCTCCAGTACAACTTA	1260
OY	421	ThrIleGluValSerAspLysAspPheLysThrLeuCysePheValGlnIleAsnValIle	440
Db	1261	ACGATAGAGGGTCTGCAGCAAAAGATTTCAGAACCTTTGTTGTGCAAATCAACCTTAT	1320
OY	441	AspIleAsnAspGlnIleProIlePheGlnLysSerAspTyzGlyAsnLeuThrLeuAla	460
Db	1321	GATATCAATGATCAGATCCCATCTTTGAAAATATAGATTATGAGAACTTGACTTCT	1380
OY	461	GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro	480
Db	1381	GAAAGCAAAACATGGGGTCCACCATCTTAACCATCCAGGCCACATGATGTATGAGCCA	1440
OY	481	PheThrGlySerSerLysIleLeuTyzIleIleLysGlyAspSerGluLysArgLeu	500
Db	1441	TTTACTGGAGATTCTTAAATCTGTATCATATACATTAAGGAGACAGAGAGGAGCTGT	1500
OY	501	GlyValAspThrAspProHisThrAsnThrGlyTyzValIleIleLysLysProLeuAsp	520
Db	1501	GGGGTGACACAGATCCCATTCACACACCGGATATGTCATATATTAAGCCCTTGAT	1560
OY	521	PheGluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal	540
Db	1561	TTTGAACACGACGCTGTTTCCAACTGTGTTCMAAGCAGAAATCTCGAGCTCTAGTG	1620
OY	541	PheGlyValLysTyzAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal	560
Db	1621	TTTGTGTGAAGTACAAATGCAGATCTTTTGGCAAGTTCACGCTTATTTGTGCAGATGTG	1680
OY	561	AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla	580
Db	1681	AATGAAGCACTCAATTTTCCCAACAGTATTCMAAGCGAAGTCAAGTGAAGGATTAAGT	1740
OY	581	IleGlyThrLysValGlyAsnValThrAlaLysAspProGluLysAspIleSerTyz	600
Db	1741	ATAGGCACTAAAGTGGGCAATGTGACTGCAAGATCCAGAAAGTCTGGACATAAGTAT	1800
OY	601	SerLeuAlaProLeuAspArgLysIleGlySerProTyzArgValGlnValAlaThr	620
Db	1801	TCACTGAGGGGAGACAAAGAGTTGGCTTAAATTTGACACGTGACTGGTGAGATCTTT	1860
OY	621	SerValAlaProLeuAspArgLysIleGlySerProTyzArgValGlnValAlaThr	640
Db	1861	AGTGTGGCTCCATTGGACAGAAACCCGGAACTCCATTCGGGTTCAAAGTGTGGCCACA	1920
OY	641	GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal	660
Db	1921	GAAGTAGGGGGGTCTTCTTAAGCTCTGTGTAGAGTTCACCTGATCTTATGAGATGTG	1980
OY	661	AsnAspAsnProProArgLeuAlaLysAspTyzThrGlyLeuPhePheCysHisProLeu	680
Db	1981	AATGCAACCCCTCCAGGCTAGCCAAAGACTACACGGGCTTGTCTTCTGCATCCCTTC	2040
OY	681	SerAlaProGlySerLeuIlePheGluAlaThrAspAspArgGlnHisLeuPheArgGly	700

Db	2041	AGTCACACGTGGAAGTCTCATTTTGGAGGCTACTGATGATGATGACGACCTTATTTCGGGGGT	2100
Qy	701	Prothiapherphreseriueuglysergiylserleuulinhsaprtppgiuvaliserlys	720
Db	2101	CCCCATTTCATTTTCCCTCGGAGGGAAGCTTACAAAACGACTGGGAAGTTTCCAAA	2160
Qy	721	lleaanglythrhisalaargyleserthrarghisrthraspphegiugluwargala	740
Db	2161	ATCATATGTTACTATGCCCCAGCTGTCTACACGACACACAGACTTTGAGGAGGCGGTAT	2220
Qy	741	ValValleuileargilleasnapglyglyarpprobleuugluyllevaliserleu	760
Db	2221	GTCCCTCTTGATCCCGACATCAATGATGGGGGGTCGGCACCCCTTGGAAGCATTTGTTCTTTA	2280
Qy	761	Provalrhrphecyssercysvalgluglysercysphearpproialglyhisglthr	780
Db	2281	CCAGTTCATTTCTTCAGTTGTGTGGAAAGAAAGTTGTTCCGGCCAGCAGGTACACAGCT	2340
Qy	781	GlylleprothrvaliglymetalaValaGlylleuleuuthrThrleuValillegly	800
Db	2341	GGGATACCCACTGTGGGCGACGTGGCATGTGCTGACACACCTTCTGTATTGGT	2400
Qy	801	llelleleualaValaValpheeliaargilleylslysaaplysglylsaspaenValglu	820
Db	2401	ATAATTTTAGAGATTGGTTATTCCGCATTAAGAAAGAGATTAAGCAAGATTAATGTTGAA	2460
Qy	821	SerlaaglinlasergluVallypProleuAargser	832
Db	2461	AGTGCTCAAGCATCTGAAGTCMAACCTCTGAGAAAGC	2496
RESULT 3			
LOCUS	AX193509	3345 bp	DNA
DEFINITION	Sequence 1076 from Patent WO0149716.	linear	PAT 15-AUG-2001
ACCESSION	AX193509		
VERSION	AX193509.1	GI:15211449	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Cranita; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J., Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.		
TITLE	Compounds for immunotherapy and diagnosis of colon cancer and methods for their use		
JOURNAL	Patent: WO 0149716-A 1076 12-JUL-2001;		
FEATURES	CORIXA CORPORATION (US)		
source	location/Qualifiers		
	1..3345		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	951 a 746 c 720 g 928 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	0	Length:	3345
Score:	4321.00	Matches:	832
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-025-380-1081 (1-832) x AX193509 (1-3345)			
Qy	1	MetlleuenglinalahisleuHiserleucysleuueMetleuTYrleuAlaThrGly	20
Db	88	ATGAACTTCAGGGCCCATCTTCACTCCCTGTGCTTCTTAAGCTTATTGGCAACTGGA	147
Qy	21	TyrGlyGlnGluGlyLysPheSerGlyProleuYspProweThrPheSerIleTyrglu	40
Db	148	TATGGCCAAAGGGGAAGTTTATGTGGACCCCTGAACCAACGACATTTTCTATTATGAA	207

QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAsnProAlaValThrPhe 60
 DB 208 GGCCAAAGACCGAATCAATATATTCAGTTTAAGCCAAATCTCTGCTGTACCTTT 267
 QY 61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluArgGlyLysLeuIleThr 80
 DB 268 GAACATACTGGGGAGACAGACAAATATTTGTATTAACGGGGAGGACTCTGTATTAAC 327
 QY 81 AsnArgAlaLeuAspArgGluThrArgSerThrIleAsnLeuGlnValAlaAlaLeuAsp 100
 DB 328 AACGAGACCTTGGACAGGAAACAAAGATCTACTCAATCTCCAGCTTGACCCCTGAC 387
 QY 101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValLysAspIleAsn 120
 DB 388 GCTAATGAAATTAAGTAGAGAGGCTCAGTCCCTATCAACATAGAAAGAGACATCAAC 447
 QY 121 AsnAsnArgProThrPheLeuGlnSerLysIleGlyGlySerValArgGlnAsnSerArg 140
 DB 448 GACATTCAGCCCAAGTTTCTCCAGTCAAGTACGAAAGCTCAGTAAAGGCAAGCTCCGC 507
 QY 141 ProGlyLysProPheLeuIleValIleAsnAlaThrAspLeuAspProAlaThrProAsn 160
 DB 508 CCAAGAAAGCCCTTCTGTATGTCAATGCCACAGACTTGATGATCCGGCCACTCCCAAT 567
 QY 161 GlyGlnLeuIleThrGlnIleValIleGlnLeuProMetIleAsnAsnValMetIlePhe 180
 DB 568 GGCAGAGCTTATTAACGATTTGATCATCAGCTTCCATGATCAACATGTCAATGATCTTT 627
 QY 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGlyGlySerGlnGluLeuAsn 200
 DB 628 CAATCAACAAACAAAGAGGAGCCATCTCTTACCCGAGAGGATCTCAGGAATGAT 687
 QY 201 ProAlaLysAsnProSerIleAsnLeuValIleSerValLysAspMetGlyGlnIleSer 220
 DB 688 CTGCTTAAGATCTCTTCTATATCTGTATCTGATCTCAAGTGAAGGACATGGAGGCAAGT 747
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGlnAsnIleThrLys 240
 DB 748 GAGATTCCTTCAAGTATCAACATCTGTGATATCATAGTGAAGAGAAATTTGGAAA 807
 QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
 DB 808 GCACCAAAACCTGTGGAGATGGGAAACCTCAATCATCTCAACCCCAACAAATCACT 867
 QY 261 GluValArgThrAspAspProGlyAlaGlnIleThrSerLeuValAspLysGlyLysLeuPro 280
 DB 868 CAGGTGGGTGATATATCCCGGTGACAAATATCTTAACTTGAAGAAAGAGCTGCCA 927
 QY 281 ArgPheProPheSerIleAspGlnGluGlyAspIleIleValThrGlnProLeuAspArg 300
 DB 928 AGATTCCTCATTTCAATTTGACCAAGAGAGATATTTAGCTGACGCCCTTGGACGA 987
 QY 301 GluGluLysAspAlaIleValIlePheIleValAlaLysAspGlyIleLysProLeu 320
 DB 988 GAAGAAAGAGATCATATGTTTATTTATGCAAGTTCGAAAGATGATACGAGAAACCACTT 1047
 QY 321 SerIleProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProProThrLys 340
 DB 1048 TCAATATCCCTGGAATTCATGTAAAGATTAATTAATATTAATCACTTACATGAT 1107
 QY 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 DB 1108 CCGTCAACAGTAAACCGATTTGAGGTCCAGAGATGAATGAAGATCAATATCCG 1167
 QY 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnIleArg 380
 DB 1168 ACCCTTAACGACATGACAGGAGTGAAGAAATATGCGCAACAGTTTCTTAACTACAG 1227
 QY 381 IleValGlnGlnIleThrProLysLeuProMetAspGlyLysLeuPheLeuIleGlnIleThrVal 400
 DB 1228 ATGTGTGAGCAAACTCCCAAACTTCCATGATGAGACTCTTCTTAATCCAAACCTTAAGCT 1287
 QY 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnIleAsnLeu 420

DB 1288 GGAATGTTACAGTTAGTAAACAGCTCTTGAAGACAGATATCTCTCAGTACAACTTA 1347
 QY 421 ThrIleGluValSerAspLysAspPheIleThrLeuCysPheValGlnIleAsnValIle 440
 DB 1348 ACGATAGAGGTGTCTGACAAAGATTTCAAGACCTTTGTTTGTGAAATCAACGTTAT 1407
 QY 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspIleGlyAsnLeuThrLeuAla 460
 DB 1408 GATATCATATCATATGATCCCATCTTGAATAATTCATATATGAAACCTGACTCTTGCT 1467
 QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 DB 1468 GAAGACAAACAACTGTGGTCCACCATCTTAACCATCAAGCCACTGATGTGATGAGCCA 1527
 QY 481 PheThrGlySerSerLysIleLeuIleThrIleIleValGlyLysAspArgGlyArgLeu 500
 DB 1528 TTTACTGGAGGATCTTAAATTTCTGTATCATATCATAAAGGAGACAGTGAAGGACCGCTG 1587
 QY 501 GlyValAspThrAspProHisThrAsnThrGlyIleValIleIleLysLysProLeuAsp 520
 DB 1588 GGGGTGACACAGATCCCATACCAACCGGATATGTCTAATTTAAAGCCTCTTGAT 1647
 QY 521 PheGluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 DB 1648 TTTGAAACAGCAGCTGTTTCCAACTGTGTGTTCAAGAGAAATCTGAGCCTTAAGT 1707
 QY 541 PheGlyValLysIleThrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 DB 1708 TTTGTGTGTAAGTACATCAATCAAGTTCTTTTGGCAAGTTCAAGCTTTTGTGACAGATG 1767
 QY 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 DB 1768 AATGAACACCTCAATTTTCCAAACGATTTCAAGCGAAAGTCAAGTGAAGTGAAGT 1827
 QY 581 IleGlyThrLysValGlyLysAsnValThrAlaLysAspProGlyLysLeuAspIleSerIle 600
 DB 1828 ATAGGCACTTAAGTGGCAATGATGATGCTGCAAGAGATCCGAAGGTCTGGACATTAACAT 1887
 QY 601 SerLeuArgGlyAspThrArgGlyIleThrLysLysIleAspHisValThrGlyLysIlePhe 620
 DB 1888 TCACCTAGGGGAGACCAAGAGGTTGGCTTAAATGACACAGCTGATGAGTCACTTT 1947
 QY 621 SerValAlaProLeuAspArgGluAlaGlySerProIleArgValGlnValAlaIleThr 640
 DB 1948 AGTGTGCTCCATTTGACACAGAGCCGGAAGTCCATATCGGTGATCAAGGTGGGCAACA 2007
 QY 641 GluValGlyLysSerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 DB 2008 GAAGTAAAGGGGCTCTCTTAAGCTGTGTGAGAGTTCCACTGATCTTATGATGATG 2067
 QY 661 AsnAspAsnProArgLeuAlaLysAspIleThrGlyLysLeuPheCysHisProLeu 680
 DB 2068 AATGACAAACCTCCCAAGGCTAGCCAAAGACTACACGGGCTTGTCTTCCATCCCTC 2127
 QY 681 SerAlaProGlySerIleLeuIlePheGluAlaThrAspAspAsnGlnIleLeuPheArgGly 700
 DB 2128 AGTGACCTGGAAGTCTATTTTCAGGCTACTGATGATGATGATGATGATGATGATGAT 2187
 QY 701 ProHisPheThrPheSerLeuGlySerGlySerIleGlnAlaAspThrProGluValSerLys 720
 DB 2188 CCCCATTTTACATTTTCCCTCCGCACTGAGAACTTACAAAGACTGGGAAGTTCCAAA 2247
 QY 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaIle 740
 DB 2248 ATCAATAGTATCATAGCCGACCTGTCTAACAGGACACAGACTTTGAGGAGAGGGGTAT 2307
 QY 741 ValValLeuIleArgIleAsnAspGlyValArgProProLeuGlnGlyIleValSerLeu 760
 DB 2308 GTGTCTTATCTCCATCATATATGAGGAGTGGGCAACCTTGAAGGCAATGTCTTTTA 2367
 QY 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnIleThr 780

Db 2368 CCAGTACATTCTGCACTGTGTGGAGAGAGTTGTTCCGGCCAGCAGGACCACT 2427
 Qy 781 GYIYLeProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuValIleGly 800
 Db 2428 GGGATACCCACTGCGGCGATGCGAGTGTGTACTGTGTGACCACTTCTGTGTGATGCT 2487
 Qy 801 ILeIleLeuAlaValIlePheIleArgIleGlyLeuAspLeuGlyLeuAspValIleGly 820
 Db 2488 ATATATTTTACGAGTGTGTATTCGCAATTAAGAGATTAAGGCAAGATATGTTGAA 2547
 Qy 821 SerAlaGlnAlaSerGluValIleProLeuArgSer 832
 Db 2548 AGTCTCAAGCATCTGAAGTCAACCTCTGAGAAC 2583

RESULT 4
 AX330272 3345 bp DNA linear PAT 09-JAN-2002
 LOCUS Sequence 781 from Patent WO0194629.
 DEFINITION AX330272
 ACCESSION AX330272
 VERSION AX330272.1 GI:18103250
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
 Horvigan, S., Soppe, D.R. and Weaver, Z.
 TITLE Cancer gene determination and therapeutic screening using signature
 gene sets
 JOURNAL Patent: WO 0194629-A 781 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES
 source 1..3345
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 951 a 746 c 720 g 928 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 3345
 Score: 4321.00 Matches: 832
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-10-025-380-1081 (1-832) x AX330272 (1-3345)

Qy 121 AspAsnArgProThrPheLeuGlnSerIleValGlySerValArgGlnAsnSerArg 140
 Db 448 GACATTCGACCCACGTTTCTCCAGTCAAAAGTGAAGGCTCAGTAAAGCAGAACTCCG 507
 Qy 141 ProGlyLeuProPheLeuThrValAlaAsnAlaThrAspLeuAspProAlaThrProAsn 160
 Db 508 CCAGAAAGCCCTTCTGTATGTCATGCAATGCAAGCAGACAGACTGATGATCCGCACTCCCAAT 567
 Qy 161 GlyIleLeuThrValGlyIleValIleGlnLeuProMetIleAsnAsnValMetIlePhe 180
 Db 568 GGCACCTTTATTAACAGATGTCATCCAGCTTCCATATGCAACATGTCATGACTTT 627
 Qy 181 GlnIleAsnAsnLeuThrGlyAlaIleSerLeuThrArgGlyIleSerGlnLeuAsn 200
 Db 628 CAGATCAACAAACAAACGGAGCCATCTCTTACCAGAGGAGATCTCAGGAATGAAAT 687
 Qy 201 ProAlaLeuAsnProSerIleAsnLeuValIleSerValIleAspMetGlyIleGlnSer 220
 Db 688 CTGTATAGATCTCTTCTATATCTGTGATCTCAGTAAAGCAGATGCGCCAGAGT 747
 Qy 221 GlnAsnSerPheSerAspThrThrSerValAspIleIleValThrGlnAsnIleTrpLeu 240
 Db 748 GAGAAATTCCTTCAGTATACCAATCTGATGATATCATATGTCAGAGAAATTTGAAA 807
 Qy 241 AlaProLeuProValGlnMetValGlnAsnSerThrAspProIleLeuIleThr 260
 Db 808 GCACCAAACTCTGTGAGATGTCGAAACCTCAACATGATCTCAACCAATCAAAATCACT 867
 Qy 261 GlnValArgIleAsnAspProGlyAlaGlnIleValIleValIleValIleValIleVal 280
 Db 868 CAGTCCGGTGAAGATGATCCGGTGCACAAATATCTTGTGTCACAAAGAAAGCTGCCA 927
 Qy 281 ArgPheProPheSerIleAspGlnGlnIleValIleValIleValIleValIleValIleVal 300
 Db 928 AGATTCCTCATTTCAATGACCAAGAGATTTTACCTGACTCAGCCCTTGACCA 987
 Qy 301 GlnGlnLeuAspAlaTrpValPheTrpAlaValAlaIleAspGlnIleValIleValIleVal 320
 Db 988 GAAAGAAAGATCATATGTTTATGATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1047
 Qy 321 SerTrpProLeuGlnIleIleValIleValIleValIleValIleValIleValIleValIleVal 340
 Db 1048 TCATATCCGCTGGAATTTCTGTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1107
 Qy 341 ProSerProAlaThrValPheGlnValGlnGlnAsnGlnArgLeuGlnIleGly 360
 Db 1108 CCGTCAACAGTAAACCGATTTAGTTCAGAGAAATGATGATGATGATGATGATGATGATGATG 1167
 Qy 361 ThrLeuThrAlaIleAspArgGlnGlnAsnThrAlaAsnSerPheLeuAsnTrpArg 380
 Db 1168 ACCCTTACTGACATGACAGGATGAGAAATTAATGCAAGCTTTCTTAACTACAG 1227
 Qy 381 IleValGlnGlnThrProLeuLeuProMetAspGlyLeuPheLeuIleGlnThrTrpAla 400
 Db 1228 ATGTGAGGCAAACTCCCAAACTCCCATGATGATGATGATGATGATGATGATGATGATGATG 1287
 Qy 401 GlyMetLeuGlnLeuAlaValGlnIleValGlnIleValGlnIleValGlnIleValGlnIleVal 420
 Db 1288 GGAATGTTTACAGTATGATTAACAGCTCTTGAAGAGCAAAATCTCTCTGATCACTTA 1347
 Qy 421 ThrIleGlnValSerAspLeuPheLeuThrLeuLeuAspPheValGlnIleValIle 440
 Db 1348 ACCATAGAGTGTCTGACAAAGATTTCAAAACCTTTGTTTGTGCAAAATCAAGTTAT 1407
 Qy 441 AspIleAsnAspGlnIlePheIlePheGlnLeuSerAspTrpGlyAsnLeuThrLeuAla 460
 Db 1408 GATATCAATGATCAAGATCCCATCTTTGAAATAATGATATGAAATCCGACCTCTTGT 1467
 Qy 461 GlnAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGlnPro 480
 Db 1468 GAAAGACCAAAACATTTGGGTCCACCATCTTAAACATCCAGGCCACTGATGCTGATGAGCCA 1527
 Qy 481 PheThrGlySerSerLeuIleLeuTrpIleIleIleValGlyAspSerGlnGlyArgLeu 500

```

Db      1528  TTTAAGGAGGAGTTCTAAATTCGTATCATATCAATAAAGGAGACAGTGGAGCGCTG 1587
QY      501  G1ValAspThrAspProHisThrAsnThyGlyTyrValIleIleIleValSLeuSProLeuAsp 520
Db      1588  GGGGTTCACACAGATCCCATACCAACACCGGATATGTCATTAATTAAGAGCCCTTGAT 1647
QY      521  PheGluThrAlaAlaValSerAsnIleValPheIleValIleValIleValIleValIleVal 540
Db      1648  TTTGAAACAGAGAGCTGTTTCCAACTGTTGTTAAACAGAAATCTGAGCCCTGATG 1707
QY      541  PheGlyValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 560
Db      1708  TTTGGTGTGAGTACAAATGCAAGTCTTTTCCCAAGTTCACGCTTATTTGACAGAGTGT 1767
QY      561  AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaIleValIleValIleValIleVal 580
Db      1768  AATGAAACCTCAATTTTCCCAACAGATTCACAAAGCAAGTCACTGAGAGATGATGCT 1827
QY      581  IleGlyThrIleValIleValIleValIleValIleValIleValIleValIleValIleVal 600
Db      1828  ATAGGCACTAAAGTGGGCAATGTGACTGCCAAGATCCAGAGGTCCTGCACTAGATCTT 1887
QY      601  SerIleuArgGlyAspThrArgGlyTyrPheIleValIleValIleValIleValIleVal 620
Db      1888  TCACTGAGGAGAGACACAGAGGTTGGCTTAAATGACACGATGAGTGGATCTTT 1947
QY      621  SerValAlaProIleuAspArgGluAlaGlySerProIleuArgGlyValIleValIleVal 640
Db      1948  AGGTGGCTCATTTGACAGAGAAAGCCGGAAGTCCATATCGGAGTACAAAGTGGCCACA 2007
QY      641  GluValGlyIleSerSerIleuSerSerValSerGluPheIleuIleuIleuIleuIleuIle 660
Db      2008  GAAGTGGGGGGGCTTCTTAAGCTCTGTGACAGATTCACCTGATCTTAAGAGATG 2067
QY      661  AsnAspAsnProProAlaGlyIleValIleValIleValIleValIleValIleValIleVal 680
Db      2068  AATGACAAACCTCCAGAGCTAGCCAGACTACACGAGCTTGTCTTCCATCCCTCC 2127
QY      681  SerAlaProGlySerIleuIlePheGluAlaThrAspAspAspGlnIleuPheArgGly 700
Db      2128  AGTGCACCTGGAAGTCTCATTTTGGAGGCTAGATGATGATGACCTTAATTTCCGGGT 2187
QY      701  ProHisPheThrPheSerIleuGlySerGlySerIleuGlnAsnAspIleuValIleVal 720
Db      2188  CCCCATTTTACATTTTCCCTCGGCAAGTGAAGCTTACAAACGACTGGGAAGTTCCAA 2247
QY      721  IleAsnGlyThrHisAlaArgIleuSerThrArgHisThrAspPheGluGluIleuValIle 740
Db      2248  ATCAATGATCATGCTGCGGACGACTGTCTACAGGACACAGACTTTGAGGAGAGGCGTAT 2307
QY      741  ValValIleuIleuArgIleuAsnAspGlyIleValArgProIleuGluIleuValIleVal 760
Db      2308  GTGCTCTGATGATCGCATCAATGATGAGGAGTGGGCGGACCTTGAAGGACTTGTCTT 2367
QY      761  ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
Db      2368  CCAAGTACATTCGAGCTTGTGGAAGAGATTTTCCGCGCAGAGAGTCAACGACT 2427
QY      781  GlyIleProThrValGlyMetAlaValGlyIleIleuIleuThrThrIleuIleuValIleGly 800
Db      2428  GGGATATCCCACTGTGGGATGAGGAGTGTGATCTGTGACACCTTCTGAGATTTGT 2487
QY      801  IleIleIleuAlaValIlePheIleArgIleIleValIleValIleValIleValIleVal 820
Db      2488  ATATATTTAGCAATGTGTGTTATCCGATTAACAAAGATTAAGCAAAAGATATGTTGAA 2547
QY      821  SerAlaGlnAlaSerGluValIleValIleValIleValIleValIleValIleValIleVal 832
Db      2548  AGTGTCAAGCACTGATGATCAAACTCTGAGAAAGC 2583

```

```

LOCUS      AX332252          3345 bp      DNA          linear      PAT 09-JAN-2002
DEFINITION Sequence 2761 from Patent WO0194629.
ACCESSION  AX332252
VERSION    AX332252.1  GI:18122886
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE  1
            Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
            Horrigan, S., Soppet, D. R. and Weaver, Z.
            Cancer gene determination and therapeutic screening using signature
            gene sets
            Patent: WO 0194629-A 2761.13-DSC-2001;
            Avalon Pharmaceuticals (US)
FEATURES
     source
         1..3345
         /organism="Homo sapiens"
         /db_xref="taxon:9606"
BASE COUNT  951 a      746 c      720 g      928 t
ORIGIN
Alignment Scores:
Pred. No.:      0      Length:      3345
Score:          4321.00      Matches:      832
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels: 0
DB:              Gaps: 0
US-10-025-380-1081 (1-832) x AX332252 (1-3345)
QY      1  MetIleleuGlnAlaHisIleuHisSerIleuCybleuIleuMetIleuIleuAlaThrGly 20
Db      88  ATGATACCTCAGGCGCCATCTTCACTCCCTGTGCTTCTTAATGTCATTTATGGCACTGGA 147
QY      21  TTYGlyGluGluGlyIleValPheSerGlyProLeuIleProMetThrPheSerIleIleValGlu 40
Db      148  TATGGCCAAAGAGGAGAAAGTTAGTGAACCCCTGAACCAATGACATTTCTATTATGAA 207
QY      41  GlyGlnIleuProSerGlnIleIlePheGlnPheIleValIleGluArgGluIleuIleuIleVal 60
Db      208  GGCCAAGAACCGAGTCAATTTATTTCCAGTTTAAGGCCAATCCCTGCTGTGACTTTT 267
QY      61  GluIleuThrGlyIleuThrAspAsnIlePheValIleGluArgGluIleuIleuIleValIle 80
Db      268  GAACATACTGGGGAGACAGACAAATATTTGATGAACGAGGAGGACTTCTGTATTTAC 327
QY      81  AsnArgAlaIleuAspArgIleuThrArgSerThrHisAsnIleuGlnValAlaIleuAsp 100
Db      328  AACAGAGCTTGGACGGGAAACAGATTTACTACCAATCTCAGGTTGACGCTGAGAC 387
QY      101  AlaAsnGlyIleIleValGluIleProValProIleThrIleGluValIleAspIleAsn 120
Db      388  GCTAATGGAATTAATAGTGAAGGTCCAGTCCCTTAATCAATGAAGTGAAGACATCAAC 447
QY      121  AspAsnArgProThrPheIleuGlnSerIleValIleValIleValIleValIleValIleVal 140
Db      448  GACAAATCGACCCAGCTTCTCCAGTCAAGTCAAGAGGCTCAGTAAAGGACGAACTCTCCG 507
QY      141  ProGlyIleProPheIleuIleValIleAsnAlaThrAspIleuAspProAlaThrProAsn 160
Db      508  CAGGAAAGCCCTTCTGTATGATGATGACAGACCTGAGATGATCGGCACTCCCAAT 567
QY      161  GlyGlnIleuIleValIleValIleGlnIleuProMetIleAsnAsnValIleMetIlePhe 180
Db      568  GGCAGAGCTTATTAACAGATTTGATCCAGCTTCCATGATCAACAGATGATGATGATGAT 627
QY      181  GlnIleAsnAsnIleThrGlyAlaIleSerIleuThrArgGluIleGlySerGlnIleuAsn 200
Db      628  CAGATCAACAAACGAGGACCATCTCTTAACCCGAGAGGAGATCTCAGAAATTTGAAT 687

```

QY 201 ProAlaIyAspAnProSerTyrrAsnLeuValIleSerValIyAspMetGlyGlnSer 220
 Db 688 CCGCTCAAGAACTCTTCTTAATCTGATCTCAAGAGAGACATGGAGCCCAAGT 747
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrPly 240
 Db 748 GAGAAATTCCTTCAGTATACCAATCTGTGGATATCATAGACAGAGAAATATTGGAAA 807
 QY 241 AlaProIyAspProValGluMetValGluAsnSerThrAspProHisProIleIleThr 260
 Db 808 GCACCAAAACCTGTGGAGATGGTGGAAAACCTCACTATCTCTCAACCCCAATCAATCACT 867
 QY 261 GluValArgTrpAsnAspProGlyAlaGlnTyrSerLeuValAspIySGluIyLeuPro 280
 Db 868 CAGGTGGGTGGATGATCCCGGTGCACATATTCCTTAGTTGACAAAGAGAGCTCCA 927
 QY 281 ArgPheProPheSerIleAspGlnGluIyAspIleTyrValThrGlnProLeuAspArg 300
 Db 928 AGATTCACATTTTCAATTGACACAGAAAGAGATATTACGTGACTCAGCCCTTGACCGA 987
 QY 301 GluGluIyAspAlaIyValPheTyrAlaValAlaIyAspGluIyTyrGlyIyProLeu 320
 Db 988 GAAAGAAAGAGATGATGTTTATGCAAGTTCAGATGCAAGATGATGATGCGGAAACCACTT 1047
 QY 321 SerTyrProLeuGluIleHisValIyValIyAspIleAsnAspAsnProProThrCys 340
 Db 1048 TCATATCCGCTGGAAATTCATGTAAGATTAAAGATTATTAATGATTAATCCACTACATGT 1107
 QY 341 ProSerProValThrValPheGluValGlnGluAsnGluIyArgLeuGlyAsnSerIleGly 360
 Db 1108 CCGTCAACAGTAACGATATTGAGGTCCAGAGAAATGACATCGGGTAAACAGATATCGGG 1167
 QY 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
 Db 1168 ACCCTTACTGCACATGACAGGATGAAGAAATATCTGCCAACATTTTCTTAACCTAAGG 1227
 QY 381 IleValGluGlnThrProIyLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 Db 1228 ATTGTGAGCAAACTCCCAAACTTCCCATGATGATGACTCTTCTTAATCCAACTTAAGCT 1287
 QY 401 GlyMetLeuGlnLeuAlaIyAspGlnSerLeuIySGlnAspThrProGlnTyrAsnLeu 420
 Db 1288 GGAATGTTACAGTAAACAGTAAACAGTCTTGAAGAGCAAGATATCTCTCACTCACTCACTTA 1347
 QY 421 ThrIleGluValSerAspIyAspPheIyThrLeuCysPheValGlnIleAsnValIle 440
 Db 1348 ACGATTAAGGTGTCTGACAAAGATTTCAGAACCTTTGTTTGTGCAAAATCAAGTTAAT 1407
 QY 441 AspIleAsnAspGlnIleProIlePheGluIySerAspTyrGlyAsnLeuThrLeuAla 460
 Db 1408 GATATCAATGATCGATCCCATCTTGTGAATAATCAAGTTATGGAACCTGACTCTTGCT 1467
 QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db 1468 GAAAGACAAACATTTGGGTCCACCATCTTAACATCCAGGCCACATGATGATGAGCA 1527
 QY 481 PheThrGlySerSerIyGlyLeuTyrHisIleIleIySGlyAspSerGluIyArgLeu 500
 Db 1528 TTTTACTGGAGATTTTAAATCTGTATCATATCAATTAAGGAGACAGTGAAGGAGCGCTG 1587
 QY 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleIySGlyProLeuAsp 520
 Db 1588 GGGGTGTGACAGATCCCATTAACCAACCCGGAATGTCAATTAATTAAGCCCTTGAT 1647
 QY 521 PheGluThrAlaAlaValSerAsnIleValPheIyValGluAsnProGluProLeuVal 540
 Db 1648 TTTTAAACAGACGCTGTTTCAACATGTTGTTCAAGAGAAATCTCGAAGCTCTAGTG 1707
 QY 541 PheGlyValIyTyrAsnAlaSerSerPheAlaIyPheThrLeuIleValThrAspVal 560
 Db 1708 TTTGTGTGAAGTACATGCAAGATTTCTTTTCCAAAGTTCAGCTTATTTGTACAGATGTG 1767
 QY 561 AsnGluIaProGlnPheSerGlnHisValPheGlnAlaIyValSerGluAspValAla 580

Db 1768 AATGAAGCACTCAATTTTCCCAACAGATATTCAGAGCGAAAGTCAAGTGAAGTATGACT 1827
 QY 581 IleGlyThrIyValIyValAsnValThrAlaIyAspProGluIyLeuAspIleSerTyr 600
 Db 1828 ATAGGACATTAAGTGGCAATGATGATGCTCCAGAGATCCAGAAAGTGTGACATTAAGCTAT 1887
 QY 601 SerLeuArgGlyAspThrArgIyTyrLeuIyIleAspHisValIleThrGlyIleIlePhe 620
 Db 1888 TCACCTAGGGGAGACACAGAGGTTGGCTTAAATTAATGACCACTGATGATGATCTTT 1947
 QY 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaIleThr 640
 Db 1948 AGTGTGCTTCATTTGACAGAGAACCGGAAGTCCATATCGGGTACAAAGTGGCGCACCA 2007
 QY 641 GluValIyGlyIySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 Db 2008 GAAATGGGGGGCTTCTTAACTCTGTGTGCAAGATTCACCTGATCTTATGATGATGTG 2067
 QY 661 AsnAspAsnProProArgIleuAlaIyAspIyTyrThrGlyLeuPheCysHisProLeu 680
 Db 2068 AATGACAACTCTCCAGGCTAGCCAGACTACACGGCTTGTCTTCTGCACTCCCTC 2127
 QY 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspGlnHisLeuPheArgIy 700
 Db 2128 AGTCACTGAGAGTCAATTTTCCAGGCTAGTATGATGATGATGATGATGATGATGATG 2187
 QY 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTrpGluValIleSerIy 720
 Db 2188 CCCCATTTTAACTTTTCCCTCCGCACTGAGAGAGCTTCAAAACACATGGGAAAGTTTCCAA 2247
 QY 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluIyValIySerIy 740
 Db 2248 ATCAATGTGATCTATGCTCCGACATCTCTACACAGCACAAGACTTTAGAGAGAGGGCTAT 2307
 QY 741 ValValLeuIleArgIleAsnAspGlyIyArgProProLeuGlnGlyIleValIleSerIy 760
 Db 2308 GTGCTGTGATCCGATCAATGATGAGGGGTCCGCACTTGAAGGCAATGTTCTTTA 2367
 QY 761 ProValThrPheCysSerCysValGluIySerCysPheArgProIleGlyHisGlnThr 780
 Db 2368 CCAATTACATTCGCAAGTGTGTGGAAGAGATGTTTCCGCGACAGGCTCCAGACT 2427
 QY 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2428 GGGATACCACCTGTGGGATGAGGATGTTGATCTGTGACCACTTCTGTGATGATGTG 2487
 QY 801 IleIleLeuAlaValAlaPheIleArgIleIyValAspIySGlyIyAspAsnValGlu 820
 Db 2488 ATATTTTACAGTGTGTTATCCGATTAAGAGATTAAGCAAAAGATTAATGTTGAA 2547
 QY 821 SerIleGlnAlaSerGluValIyProLeuArgSer 832
 Db 2548 AGTGTCCAGCATCTGAAGTCAAACTCTGAGAGAGC 2583
 RESULT 6
 AX410642 3345 bp DNA linear PAT 14-JUN-2002
 LOCUS Definition Sequence 3289 from Patent WO0229103.
 ACCESSION AX410642
 VERSION AX410642.1 GI:21443347
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
 TITLE Gene expression profiles in liver cancer
 JOURNAL Patent: WO 0229103-A 3289 11-APR-2002;
 GENE LOGIC INC (US)
 FEATURES location/Qualifiers
 source 1..3345

/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. U07963"

BASE COUNT 951 a 746 c 720 g 928 t

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	3345
Score:	4321.00	Matches:	832
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-025-380-1081 (1-832) x AX410642 (1-3345)

QY 1 MetIleuGlnAlaHisIleuHisSerLeuCySleuLeuMetLeuTYrLeuAlaThrGly 20
DB 88 ATGATACCTTCAGGCCCACTTCCTCCCTGTCCTTATGCTTTATTTGGCAACTGGA 147
QY 21 TYrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTYrGlu 40
DB 148 TATGGCCAAAGGGGAAAGTTTATGAGACCCCTGAAACCCATGACATTTTCTATTATGAA 207
QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAsnProProAlaValThrPhe 60
DB 208 GGCCAAAGACCGAATCAATTAATTCAGTTTAAAGCCAAATCCTCGCTGTGACTTTT 267
QY 61 GlnLeuThrGlyGlyThrAspAsnIlePheValIleGluArgGlyGlyLeuLeuTYrTYr 80
DB 268 GAATTAATCGGGAGACAGACAAACATATTTGTATTAACGGAGGAGGACTTGTATATAC 327
QY 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
DB 328 AACAGACCTTGGACAGAGAAACAAGATCTACTACATCTCCAGCTTGCAGCCCTGAC 387
QY 101 AlaAsnGlyIleIleValGlnGlyProValProIleThrIleGluValLysAspIleAsn 120
DB 388 GCTAATGGAATTAATGAGAGGGTCCAGTCCCTATCAACATAGAAAGTGAAGCATCAAC 447
QY 121 AsnAsnArgProThrPheLeuGlnSerIleTYrGlyGlySerValArgGlnAsnSerArg 140
DB 448 GACAAATGACCCAGTTCCTCCAGTCAAGTACGAAAGCTCAGTAAGGCAAGCTTCGC 507
QY 141 ProGlyLysProPheLeuTYrValAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
DB 508 CCAAGAAAGCCCTTCTTGTATGTAATGCAAGACCTGGATGATCCGGCCACTCCCAAT 567
QY 161 GlyGlnLeuTYrTYrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTYrPhe 180
DB 568 GGCACAGCTTATTAACAGATGTCATCAGCTCCATGATCAACATGTCATGTACTTT 627
QY 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGlyGlySerGlnGluLeuAsn 200
DB 628 CAGATCAACAACAAGAGGAGGACATCTCTTACCAGAGAGGAGATCTCAGGAATTAAT 687
QY 201 ProAlaLysAsnProSerTYrAsnLeuValIleSerValLysAspMetGlyGlnGlnSer 220
DB 688 CCGCTAGAAATCTCTTAATATCTGTATGATCAATGAAAGCAATGAGGACCAAGT 747
QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTYrLys 240
DB 748 GAGAAATCTTCACTGATACCAATCTGTGGAATATCAATGACAGAAATATTGGA 807
QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
DB 808 GCACCAAAACCTGTGGAGATGGTGAAGAACTCAACTATCTCCACCCCAATCAATAC 867
QY 261 GluValArgTrpAsnAspProGlyAlaGlnTYrSerLeuValAspLysGlyLysLeuPro 280
DB 868 CAGGTGGCTGGATGATCCCGGTGACAAATATCTTATGTAACAAGAGAGCTGCCA 927
QY 281 ArgPheProPheSerIleAspGlnGlyAspIleTYrValThrGlnProLeuAspArg 300

DB 928 AGATTCCTCATTTCAATTCAGCAGAGAGATTTATCGTCACTGACCCCTTGACACGA 987
QY 301 GluGluLysAspAlaTYrValPheTYrAlaValAlaLysAspGluTYrGlyLysProLeu 320
DB 988 GAAAGAAAGATCATATGTTTATTTATGACAGTTGCAAGATAGTACGAAACCACTT 1047
QY 321 SerTYrProLeuGluIleHisValLysValLysAspIleAsnAspAsnProProThrCys 340
DB 1048 TCATATCCGTGGAAATTCATGTAAGTTAAAGTAAATTAATATATCACTACATG 1107
QY 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
DB 1108 CCGTCAACAGTAAACCGATTTGAGTCCAGAGAAATGAAGCATCGGGTACAGATCGGG 1167
QY 361 ThrLeuThrAlaHisAspArgGluGluAsnThrAlaAsnSerPheLeuAsnTYrArg 380
DB 1168 ACCCTTACTGCACATGACAGAGATGAAGAAATACGCCAAAGTTTCTAACTACAG 1227
QY 381 IleValGlnGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTYrAla 400
DB 1228 ATTGTGAGCAAACTCCCAACTTCCCATGATGACTCTTCTTAATCCAACTATGCT 1287
QY 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTYrAsnLeu 420
DB 1288 GGAATGTACAGTTAGTAAACAGTCTTGAAAGAACAGATATCTCCTCAGTCAACTTA 1347
QY 421 ThrIleGluValSerAspLysAspPheLysThrLeuCyPheValGlnIleAsnValIle 440
DB 1348 ACATATAGAGGTGTCTCAACAAGATTCAGAACCTTGTGTTGTGCAAACTGATTA 1407
QY 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTYrGlyAsnLeuThrLeuAla 460
DB 1408 GATATCAATCATCATGATCCCATCTTGTGAATAATCGATTAATGAACCTGACTTGTCT 1467
QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
DB 1468 GAAGACAAACAATGGGTCCACATCTTAACCATCAAGCCAGCTGATGAGCCA 1527
QY 481 PheThrGlySerSerLysIleLeuTYrHisIleIleLysGlyAspSerGlyValGluLeu 500
DB 1528 TTTACTGGGAGTCTAAATCTGTATCATATCAATTAAGGGAACAGTGGAGCAGCTG 1587
QY 501 GlyValAspThrAspProHisThrAsnThrGlyTYrValIleIleLysLysProLeuAsp 520
DB 1588 GGAGTTGACACACATCCCATCCATCAACACCGATATGTCATATTAAGGCTCTTGAT 1647
QY 521 PheGluThrAlaAlaLysSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
DB 1648 TTTGAAACAGCAGCTGTTTCCAACATTTGTGTTCAAGCAAGAAATCCTGAGCTTATG 1707
QY 541 PheGlyValLysTYrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
DB 1708 TTTGGTGTGAAGTACATCAAGTCTTTTGGCCAACTTCAACGCTTATTTGACAGATG 1767
QY 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
DB 1768 AATGAAGACCTCAATTTTCCAAACGATTTCAAGCAAGAACTCAGTGAAGTATGCT 1827
QY 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGlyGlyLeuAspIleSerTYr 600
DB 1828 ATAGGCACTTAAGTGGGCAATGATGATCGCAAGAGATCCGAAGGTCTGGACATTAAC 1887
QY 601 SerLeuArgGlyAspThrArgGlyTYrPheLysIleAspHisValThrGlyGluIlePhe 620
DB 1888 TCACCTAGGGGAGACCAAGAGGTTGGCTTAAATTAATGACACAGTGACTGGTGAATCT 1947
QY 621 SerValAlaProLeuAspArgGluAlaGlySerProTYrArgValGlnValAlaThr 640
DB 1948 AGTGTGCTTCATTTGACAGAGAGCCGGAAGTCCATATCGGGTACAAAGTGGGCA 2007
QY 641 GluValGlyGlySerSerLeuSerValSerGluPheHisLeuIleLeuMetAspVal 660

Db 2008 GAAGTAGGGGGCTTCTTAAGCTCTGTGTGACAGATTCCACCTGATCTTATGAGATGTG 2067
Qy 661 AsnaapshenProProAlaGluAlaAlaGlySerThrGlyLeuPhePheCysHisProLeu 680
Db 2068 AATACACACCTCCAGAGCTTACCAAGACATACAGGGCTGTCTTCTGACATCCCTC 2127
Qy 681 SerAlaProGlySerLeuLeuPheGluAlaThrAspAspGlnHisLeuPheArgGly 700
Db 2128 AGTCACCTGGAGATCTCATTTTTCGAGGCTACTATGATGATGACACTTTTTCGGGGCT 2187
Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTrpGluValSerLys 720
Db 2188 CCCATTTTTCATTTTCCCTCGGACGTGGAGAGCTTACAAAGCACTGGAGAGATTTCACAA 2247
Qy 721 IleAsnGlyThrHisAlaAlaGlyLeuSerThrArgHisThrAspPheGluValArgAlaTyr 740
Db 2248 ATCAATGATGATCACTGCGCCGACTGTCTACACGACACACAGCTTTGGAGAGGGCCGTAT 2307
Qy 741 ValValLeuLeuIleArgIleAsnAspGlyGlyValArgProProLeuGluGlyIleValSerLeu 760
Db 2308 GTCTCTTGTATCCGATCATCATGATGGGGGTGGCCACCTTGGAGAGGATTTGTTCTTTA 2367
Qy 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
Db 2368 CCAGTTACATCTCGACGTTGTGTGGAGAGAGTTGTTCCGGCAGCAGGTCAACAGACT 2427
Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
Db 2428 GGGATACCCACTGCGGGGATGAGCAGTTGTATGCTGACACACCTTCTGTGATTTGT 2487
Qy 801 IleIleLeuAlaValAlaPheIleArgIleLysIleAspLysGlyLysAspAsnValGlu 820
Db 2488 ATAAATTTTAGCAGATGTTGTGTTATTCGCAATTAAGAGATTAAGCAAGATTAATTTGAA 2547
Qy 821 SerAlaGlnAlaSerGluValLysProLeuArgSer 832
Db 2548 AGTCTCAAGCATCTGAAGTCAAACCTCTGAGAGGC 2583
RESULT 7
HSU07969 3345 bp mRNA linear PRI 15-JUN-1994
LOCUS Human intestinal peptide-associated transporter HPT-1 mRNA,
DEFINITION complete cds.
ACCESSION U07969
VERSION U07969.1 GI:483391
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3345)
AUTHORS Dantzig,A.H., Hoskins,J., Tabas,L.B., Bright,S., Shepard,R.L.,
Jenkins,I.L., Duckworth,D.C., Sportsman,R., Mackensen,D.,
Roestek,P.R. and Skatrud,P.L.
TITLE Association of intestinal peptide transport with a protein related
to the cadherin superfamily
JOURNAL Science 264 (5157), 430-433 (1994)
MEDLINE 94204643
PUBMED 8153632
REFERENCE 2 (bases 1 to 3345)
AUTHORS Roestek,P.R.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-1994) Paul R. Roestek Jr., Lilly Research
Laboratories, Eli Lilly and Co., Lilly Corporate Center,
Indianapolis, IN 46285, USA
FEATURES
source 1..3345
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Caco-2"
/cell_type="adenocarcinoma"
/tissue_type="colon"
88..2586
CDS

/codon_start=1
/product="intestinal peptide-associated transporter HPT-1"
/protein_id="AAA19021.1"
/db_xref="GI:483392"
/translation="MTLQAHLSLCLMLYATGCGQEGFSGPLKPMFTSYEGEP
SGLIFQKAMPAYVELTGETDNI FVIEREGLLYYNRALDRSTRNLHQLVALDAN
GLIVEGPAVITTEEDKINDNRPFELQSKRFGSROMSRGPFLVYNADLDPAEPN
GOLYQVIVQLPMINVMYFQINNKTKGALSIFREGSOELNPNANPSYLVISVMDG
OSNBSRSDTSVDIIVTENIWKAPREVENVNSDHPKIKITQVRNNDGQASLVKD
EKLPREPSITDDEGDIYVQPDREKDAVPAVAKDXEGRPLSYLETHVYKDN
DNEPTSPSPVTFEVEENERLGNISITLITLTVESDKPFLTCFQVINDIINDIPIE
LELIQTAYAMDLQAKSLKODTPQNLITLTVESDKPFLTCFQVINDIINDIPIE
KSDYGNLTLEADENISITLITLTVESDKPFLTCFQVINDIINDIPIE
NNGVYIKKPLDEPTAASNIYKAPNEPELVGVYKASSEFAKFLIYDVVEAQF
SOHVPAKXSEVDAIGTQVNTYAKDPEGLDISYLRGDRGMLKIDHTGEIFSVAP
LDREASSPVRQVYATVEVGSSLSYSYSESHLIMVDYNDNPPRLAKDYTGIFCHPLSA
PGLSILEATVDHDLFRGHFTFSLSSGSLQNDMEVSKNGTARLSTHTDEEAY
VVLIRINDGRPEPLGIVSLPVTFSCVSGSRPFGHQTGLPTVMAVGILTLTLV
IGIILVVFIRIKDKGKDNVSAQSEVYKPLRS"
BASE COUNT 951 a 746 c 720 g 928 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3345
Score: 4321.00 Matches: 832
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-025-380-1081 (1-832) x HSU07969 (1-3345)
Qy 1 MetIleleuGlnAlaHisleuHisserLeuCysleuLeuMetLeuTyrleuAlaThrGly 20
Db 88 ATATACCTTAGGCCCATCTTCACTCCCTGTGCTTTTATGATTTATTTGGCAACTGGA 147
Qy 21 TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGlu 40
Db 148 TATGGCCAGAGGGGAAAGTTTACTGGACCCCTGAACCCATGACATTTCTATTATGAA 207
Qy 41 GlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAsnProProAlaValThrPhe 60
Db 208 GGCCAGAGAACCGAGTCAATATATTCAGTTTAAGGCCATCTCTGCTGACATTTT 267
Qy 61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluArgGlyLeuLeuTyrTyr 80
Db 268 GAACCTAACCTGGAGACAGACACATATTTGTATGAAAGGAGGACCTTCTGTATTAC 327
Qy 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
Db 328 AACAGAGCCTTGGACAGGAAACAGATCTACTCACATCTCCAGGTTGCAGCCCTGGAC 387
Qy 101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValLysAspIleAsn 120
Db 388 GCTAATGAAATTATAGGAGGGGTCCAGTCCCTATACCCATAGAGTGAAGGACATCAAC 447
Qy 121 AspAsnArgProThrPheLeuGlnSerLeuTyrGluGlySerValArgGlnAsnSerArg 140
Db 448 GACATATGACCCAGCTTCTCCAGTCAAGATGAGAGGCTCAGTAAAGGCACTCTCCG 507
Qy 141 ProGlyLysProPheLeuTyrValAlaAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
Db 508 CCAGAGAACCTTCTTGTATGTCATACAGACAGACCTGATGATCCGGCAGCTCCCAAT 567
Qy 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
Db 568 GGCCAGCTTATTAACGATGTCATCCAGCTTCCCATGATCAACAATGTCATGATCTTT 627
Qy 181 GlnIleAsnLeuTyrThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
Db 628 CAGTTTCAACACAAAGGAGGACATCTCTTACCCGAGAGGGATCTCGAGAAATGTAAT 687
Qy 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnSer 220

Db 688 CCTCTAGAGATCCCTTCCATTAATCTGTGATCTCAGTAGAGACATGGGAGGCGCAGAGT 7447
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTrpLys 240
 Db 748 GAGAAATCTTCACTGATACCAATCTGTGATATCTAGTACAGAGAAATATTGGAAA 807
 QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
 Db 808 GCACCAAAACCTGTGAGATGTGGAAAACCACTGATCTCACTCCCATCAAAATCACT 867
 QY 261 GluValArgTrpAsnAspProGluValArgIleYserIleuValAspLysGluLysLeuPro 280
 Db 868 CAGGTGGGTGGATGATCCCGGACCAATATTCCTTAGTTGACAAAGAGAACCTCCCA 927
 QY 281 ArgPheProPheSerIleAspGluGluValAspIleYserValThrGluProLeuAspArg 300
 Db 928 AGATTCCCATTTTCAATTGACACAGAAAGAGATATTTACGTGACTCAGCCCTTGGACCGA 987
 QY 301 GluGluLysAspAlaYrValPheYrAlaValAlaLysAspGluYrGluYrLysProLeu 320
 Db 988 GAAGAAAGAGATGATATGTTTATTCAGTTGCCAAAGATGAGTACGGAAAACACTT 1047
 QY 321 SerYrTrpLeuGluIleHisValLysValLysAspIleAsnAspAsnProTrpThrCys 340
 Db 1048 TCATATCCGCTGGAAATTCATGTAAAGATTAAAGATTATATGATATCCACCTACACTGT 1107
 QY 341 ProSerProValThrValPheGluValGluGluAsnGluArgLeuGluYrAsnSerIleGluY 360
 Db 1108 CCGTACACAGTAAACCGATTGTAGGTCCAGAGATACACACTGGGTAAACAGATCGGG 1167
 QY 361 ThrIleuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnYrArg 380
 Db 1168 ACCCTTACTGACATGACAGGAGTGAAGAAATATCTGCCAAACGTTTCTTAATACAGG 1227
 QY 381 IleValGluGluThrProLysLeuProMetAspGluYrLeuPheLeuIleGluThrYrAla 400
 Db 1228 ATGTGTGAGCAAACTCCCAAACTTCCATGATGATGATCTCTCCAAATCCAAACCTAATGT 1287
 QY 401 GluMetLeuGluLeuAlaLysGluSerLeuLysGluAsnThrProGluYrAsnLeu 420
 Db 1288 GGAATGTTCAGTTAGCTTAACACTCTTGAAGAGAAATACCTCTCGTACGTAACCTTA 1347
 QY 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGluIleAsnValIle 440
 Db 1348 ACGATGAGGTGTCTGACAAAGATTTCAAGACCTTTGTTTGGCAATCAACGTTATT 1407
 QY 441 AspIleAsnAspGluIleProIlePheGluLysSerAspYrGluYrAsnIleuThrLeuAla 460
 Db 1408 GATATCAATGATCAGATCCCATCTTGTGAATAATCAGATTATGGAACCTGACTCTTGT 1467
 QY 461 GluAspThrAsnIleGluYserThrIleLeuThrIleGluAlaThrAspAlaAspGluPro 480
 Db 1468 GAAAGACAAACATGGGTCCACCATCTTAACATCCAGGCCACTGATGTGATGAGCCA 1527
 QY 481 PheThrGlySerSerLysIleLeuYrHisIleIleLysGluYrAspSerGluGluYrArgLeu 500
 Db 1528 TTTACTGGGAGTTCTAAATTTCTGTATCATATTAAGGAGACAGTGGAGGAGCCTG 1587
 QY 501 GlyValAspThrAspProHisThrAsnThrGlyYrValIleIleLysLysProLeuAsp 520
 Db 1588 GGGGTGTGACACATCCCATACCAACACCGGATATGTCTAATTTAAAGGCTCTTGAT 1647
 QY 521 PheGluThrThrAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 Db 1648 TTTGAAACGACACTGTTTCCAACTTGTGTCAAGCAAAAATCCTGAGCCTTAGTG 1707
 QY 541 PheGluValLysYrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 Db 1708 TTTGGTGTGAAGTACATGCAAGTCTTTTGGCCAGTTCACGCTTATTTGACAGATGTG 1767
 QY 561 AsnGluAlaProGluPheSerGluHisValPheGluAlaLysValSerGluAspValAla 580

Db 1768 AATGAAGACCTCAATTTTCCCAACGTAATTCACGAAAGTACGTAGTAGTACT 1827
 QY 581 IleGluYrThrLysValGluAsnValThrAlaLysAspProGluGluYrLeuAspIleSerYr 600
 Db 1828 ATAGGCACTAAAGTGGCAATGTGACTGCCAAGGATCCAAAGAGTCTGGACATAGCTAT 1887
 QY 601 SerLeuArgGluYrAspThrArgGluYrTrpLeuLysIleAspHisValThrGluIleIlePhe 620
 Db 1888 TCACGTAGGGAGACACAAAGAGTGGCTTAAATTAATGACACAGTGTGATGATCTTT 1947
 QY 621 SerValAlaProLeuAspArgGluAlaGluYserProYrArgValGluValAlaThr 640
 Db 1948 AGTGTGGCTCCATTGACAGAGAGCCGGAAGTCCATATCGGTACAAAGTGGGCCACA 2007
 QY 641 GluValGluYrGlySerSerLeuSerSerValIleSerGluPheHisIleuIleLeuMetAspVal 660
 Db 2008 GAAAGTAAAGGGGGCTTCTTCACTTCACTGTGTCAAGATTCACCTGATCTTATGAGATGTG 2067
 QY 661 AsnAspAsnProProArgLeuAlaLysAspYrThrGluYrLeuPhePheCysHisProLeu 680
 Db 2068 AATGACAAACCTCCCAAGGCTAGCCAAAGACTACACGGGCTGTCTTCTGCCATCCCTTC 2127
 QY 681 SerAlaProGluYserLeuIlePheGluAlaThrAspAspAspGluHisLeuPheArgGly 700
 Db 2128 AGTGCACCTGGAGAGTCTCATTTTCGAGGCTACTGATGATGATCAGCATTAATTCGGGGT 2187
 QY 701 ProHisPheThrPheSerLeuGluYserGluYserLeuGluAsnAspTrpGluValSerLys 720
 Db 2188 CCCCATTTTACATTTTCCCTCGGACGTGAAAGCTTCAAAACACATGGGAAGTTTCCAA 2247
 QY 721 IleAsnGluYrThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaYr 740
 Db 2248 ATCAATGTACTCATATCCCGACGTCTTACCAAGCACAACAGACTTTAAGAGAGGGCTAT 2307
 QY 741 ValValLeuIleArgIleAsnAspGluYrArgProProLeuGluGluIleValSerLeu 760
 Db 2308 GTGCTCTTATCCGCAATCATATGATGGGGTCCGCCACCTTGGAAAGCAATGTTCTTTA 2367
 QY 761 ProValThrPheCysSerCysValGluYserCysPheArgProAlaGluHisGluThr 780
 Db 2368 CCAATTAACATCTGCAAGTGTGTGGAAGAGATGTTTCCGCGACAGTCCACGACT 2427
 QY 781 GlyIleProThrValGluYrMetAlaValGluIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2428 GGGATACCACTAGTGGGCAATGGCAGTGGTATATCTGTACCACTTCTGGATTTGCT 2487
 QY 801 IleIleLeuAlaValAlaPheIleArgIleLysLysAspLysGluYrLysAspAsnValGlu 820
 Db 2488 ATTAATTTTACAGTTGTGTATCCGCAATTAAGAGATTAAGGCAAAAGATATATGTGA 2547
 QY 821 SerAlaGluAlaSerGluValLysProLeuArgSer 832
 Db 2548 AGTGTCAAGCAATCTGAAAGTCAAACTCTGAAGAAC 2583
 RESULT 8
 AX330312 3654 bp DNA linear PAT 09-JUN-2002
 LOCUS Sequence 821 from Patent WO0194629.
 DEFINITION AX330312
 ACCESSION AX330312
 VERSION AX330312.1 GI:18103290
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1
 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
 Horrigan, S., Soppet, D.R. and Weaver, Z.
 TITLE Cancer gene determination and therapeutic screening using signature
 gene sets
 JOURNAL Patent: WO 0194629-A 821 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES Location/Qualifiers

source 1..3654
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1052 a 804 c 777 g 1021 t
ORIGIN

Alignment Scores:
Prod. No.: 1,996-317 Length: 3654
Score: 4303.00 Matches: 828
Percent Similarity: 99.76% Conservative: 2
Best Local Similarity: 99.52% Mismatches: 2
Query Match: 99.58% Indels: 0
DB: Gaps: 6

US-10-025-380-1081 (1-832) x AX330312 (1-3654)

QY 1 MetIleuGlnAlaHisLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
DB ATGATCTTCAGGCGCCATCTTCACTCCCTGTGTCTTATGCTTTATTTGGCACTGGA 152

QY 21 TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGlu 40
DB TATGGCCAGAGGGGAAATTAGTGAACCCCTGAACCCATGACATTTCTATTATGA 212

QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAsnProAlaValThrPhe 60
DB GGCCAGAACCCGATCAATATTTATTCAGTTTAAGCCCATCTCTGCTGATCTTT 272

QY 61 GluLeuThrGlyGlnThrAspAsnIlePheValIleGluArgGluGlyLeuLeuTyr 80
DB GAACTAATCGGGAGACAGACACATATTTGTATGATGAACGGGAGGAGCTTGTATAC 332

QY 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaIleAsp 100
DB AACAGACCTTGGACAGGAAACAAAGATCTACATCTCCAGTTGACGCTTGAC 392

QY 101 AlaAsnGlyIleIleValGlnGlyProValProIleThrIleGluValLysAspIleAsn 120
DB GCTATATGAATTAATGAGGAGGCTCCAGTCTTACCATTAAGTGAAGACATCAAC 452

QY 121 AspAsnArgProThrPheLeuGlnSerLysTyrGlnGlySerValArgGlnAsnSerArg 140
DB GACATGACACCCATCTTCTCCAGTCAAAAGTACAAAGGCTCAGTAAGGACATCTGC 512

QY 141 ProGlyLysProPheLeuTyrValAlaAsnAlaThrAspLeuAspProAlaThrProAsn 160
DB CCAGAAAGCCCTTCTGTATGATCAATGCCAGACCTGATGATCCGGCCATCCCAAT 572

QY 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
DB GGCCAGCTTTATTCAGATTTGTATCCAGCTTCCATGATCAACATGTCTATCTTT 632

QY 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
DB CAGATCAACAAACAAACGAGGACCATCTCTTACCCGAGAGGATCTCAGAAATTAAT 692

QY 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnSer 220
DB CTCTCTAGATCTCTTCTATTAATCTGTGATCTCAGTGAAGGACATGGGAGGCAAGT 752

QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrLys 240
DB GAGAAATCTTCAATGATCAACATCTGTGATATCATAGTGAAGAAATTTTGGAAA 812

QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
DB GCACCAAAACCTGTGAGATGTGGAAACCTCACTGATCTCAACCCCATCAAAATCACT 872

QY 261 GlnValArgTyrPheAspProGlyAlaGlnTyrSerLeuValAspLysGluLysLeuPro 280
DB CAGGTGGGTGGAATGATCCGGGTGCACAATATTTCTTAAGTTGACAAAGAACTGCA 932

QY 281 ArgPheProPheSerIleAspGlnGluLysAspIleTyrValThrGlnProLeuAspArg 300

DB 933 AGATTCCTCAATTTCAATGACCGAAGAGATATTAGTGACTCAGCCCTTGGACCGA 992
QY 301 GlnGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu 320
DB GAGAAAGAGATGATATGTTTTTATGATGATGCAAGGATGATGAGAAACACATTT 1052

QY 321 SerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProThrCys 340
DB TCATATCCCTGGAATTCATGTAAGTTAAAGATTAATATATATATCATCACTCATATG 1112

QY 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlnLysSerIleGly 360
DB CCCTCACACAGTAACCGTATTTGAGTCCAGGAGAAAGAACGACTGGTATACATATCGGG 1172

QY 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
DB ACCCTTAATGACATATACAGAGGATGAAGAAATACGCAACGTTTCTTAACTACAGG 1232

QY 381 IleValGlnGluThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
DB ATTGTGAGCAAACTCCCAACTTCCCATGATGAGACTCTTCTTATCCAACTATGCT 1292

QY 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyrAsnLeu 420
DB GGAATGTTACGTTAGCTAAACAGTCTTGAAGAACCAAGATATCTCTCAGTCAAACTTA 1352

QY 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
DB ACATATGAGGTGTCTCAACAAAGTTTCAGAACCTTGTGTGTGCAAACTCAAGTATAT 1412

QY 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAla 460
DB GATATCAATGATACAGCCCATCTTGAATAATCAGATTATGGAACCTGATCTTGCT 1472

QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
DB GAAAGACAAACATGTGGTCCACCATCTTAACATCAGGCGCCATGCTGATGAGCCA 1532

QY 481 PheThrGlySerSerLysIleLeuTyrHisIleIleGlyLysAspSerGluArgLeu 500
DB TTTTACGGGAGTTCTAAATTTGTATCATATCAATTAAGGAAAGCATGAGGACGCTG 1592

QY 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
DB GGAGTTGACACATATCCCATACCAACCGGATATGTCTAATTAATAAGCCTCTGAT 1652

QY 521 PheGluThrAlaAlaLysSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
DB TTTGAAACAGCAGCTGTTTCCACATGTTGTCAAGCAGAAATCCTGAGCCTTATGTG 1712

QY 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValIThrAspVal 560
DB TTTGGTGTGAAGTACATGACAGTCTTTTGGCAAGTTCAGCTTATTTGGACAGATGTG 1772

QY 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
DB AATGAGCAGCTCAATTTTCCCAACAGTATTTCCAGCGAAATGATGAGTGAATGCT 1832

QY 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTyr 600
DB ATAGGACCTAAAGTGGCAATGTGATGCTGCAGAGATCCAAAGTGTGAGCATTAAGCTAT 1892

QY 601 SerLeuArgGlyAspThrArgGlyTyrPheLysIleAspHisValThrGlyGluIlePhe 620
DB TCAGTGGGGAACACCAAGAGTGGCTTAAATTAATGACCACTGATGTGTGAGATCTTT 1952

QY 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaThr 640
DB AGTGTGCTCATTTGACAGAGAACCGGAGTCCATATGTGGGTACAAAGTGTGGCCACA 2012

QY 641 GluValGlyLysSerSerLeuSerValSerGluPheHisLeuIleLeuMetAspVal 660

Db 1173 ACCCTTACTCAGACAGGAGTGAAGAAATCTCCACAGCTTTCTTAACTACAGG 1232
 Qy 381 ILeValGluGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyra 400
 Db 1233 ATGTGGAGCAACTCCCAACCTTCCATGATGATGATCTTCTTAACTCAACCTATGCT 1292
 Qy 401 GlyMetLeuGlnLeuAlaValGlnSerLeuLysLeuGlnPheThrProGlnTyraMetLeu 420
 Db 1293 GGAATGTTCAGTAAAGTAAACAGTCTTGAAGAGCAAGTACTCTCTCACTCACTTA 1352
 Qy 421 ThrIleGlnValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
 Db 1353 ACCGATAGAGGTCTGTGCAAAAGATTTCAAGACCTTTGTTTGCAATCAAGCTTATT 1412
 Qy 441 AspIleAsnAspGlnIleProIlePheGlnLysSerAspTyrglyAsnLeuThrLeuAla 460
 Db 1413 GATATCAATGATGACCCCATCTTGAATAATCAATATGAGAACTGACTCTTGCT 1472
 Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db 1473 GAAGACACAAACATTTGGGTCCACATCTTAACATCCAGGCACTGATGCTGATGAGCCA 1532
 Qy 481 PheThrGlySerSerLysIleLeuTyriIleIleLysGlyAspSerGlnGlyArgLeu 500
 Db 1533 TTTACTGGAGGTTTAAATTTCTGTATCATCTTAAAGGAGACAGTGGAGGAGCGCTG 1592
 Qy 501 GlyValAspThrAspProHisThrAsnThrGlyTyraIleIleLysLysProLeuAsp 520
 Db 1593 GGGGTGACACAGATCCCAATCCACACCGGATATGTCATTAATTAAGCCCTTGAT 1652
 Qy 521 PheGluThrAlaAlaValSerAsnIleValPheValAlaGluAsnProGluProLeuVal 540
 Db 1653 TTTGAAACAGACGCTGTTCCACATGTTGTTCAAGAGCAAAATCTGAGCCCTGATG 1712
 Qy 541 PheGlyValLysTyraAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 Db 1713 TTTGCTGTGAAGTCAATGCAATGCTTTTCCCAAGTTCACGCTTATTTGTCAGAGTGTG 1772
 Qy 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 Db 1773 AATGACACACCTCAATTTTCCCAACGATTTCCAGAGCAAAAGTCAAGATGATGATCT 1832
 Qy 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTy 600
 Db 1833 ATAGGCACTAAAGGGGCAATGTGATGCTCCAGAGTCCAGAGGCTCGAGCATAGCAT 1892
 Qy 601 SerLeuArgGlyAspThrArgGlyTyrLeuLysIleAspHisValThrGlyGluIlePhe 620
 Db 1893 TCACTGAGGGAGACACAAAGGCTTGCTTAAATTTGACACGAGCTGGTGAATCTTT 1952
 Qy 621 SerValAlaProLeuAspArgGluAlaGlySerProTyraArgValGlnValAlaThr 640
 Db 1953 AGTGTGCTCATTTGACAGAGAACCGGAGTCCATATCGGTGATCAAGTGTGTGCCACA 2012
 Qy 641 GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleMetAspVal 660
 Db 2013 GAAATAGGGGGGCTTCTTCAAGCTGTGTGTCAAGTTCACATGATCTTATGAGATGTG 2072
 Qy 661 AsnAspAsnProProArgLeuAlaLysAspTyraThrGlyLeuPhePheCysHisProLeu 680
 Db 2073 TATATCAACCTCCCAAGCTAGGAGGACATACAGGGCTTGTCTTGTGCAATCCCTC 2132
 Qy 681 SerAlaProGlySerLeuIlePheGlnAlaThrAspAspAspGlnHisLeuPheArgGly 700
 Db 2133 AGTCACCTGGAAGTCTATTTGAGGCTACTGATGATGATGACACTTATTTGCGGGT 2192
 Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerLys 720
 Db 2193 CCCCATTTTACATTTTCCCTGGGAGTGGAGCTTACAAACGACTGGGAAGTTTCCAAA 2252
 Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluAlaValTy 740

Db 2253 ATCAATGTACTCAGTCCCGCACTGTCTACCAAGACACAGAGTTTGAGGAGGAGTAT 2312
 Qy 741 ValValLeuIleArgIleAsnAspGlyValArgProProLeuGlnGlyIleValSerLeu 760
 Db 2313 GTCTGTTATCCGATCATATGAGGGGTGGCCACCTTGGAGAGCATTTGTTCTTAA 2372
 Qy 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaValHisGlnThr 780
 Db 2373 CCAATTCATTCGTCAGTGTGTGTGAGAGAGATTTTCCGGCAGCAGTACCAAGACT 2432
 Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuValIleGly 800
 Db 2433 GGGATACCACTGTGGGAGTGGAGTGGATATCTGTCACCACTCTGTGTGATTTGT 2492
 Qy 801 IleIleLeuAlaValAlaPheIleArgIleValLysAspLysGlyLysAspAsnValGlu 820
 Db 2493 ATATATTTAGCAGTGTGTGTTATCCCAATTAAGAGATTAAGCAAGATTAATTTGA 2552
 Qy 821 SerAlaGlnAlaSerGluValLysProLeuArgSer 832
 Db 2553 AGTCTCAAGCATCTGAAGTCAAACTCTGAGAGAGC 2588
 RESULT 10
 AX411163
 LOCUS AX411163 3654 bp DNA linear PAT 14-JUN-2002
 DEFINITION Sequence 3810 from Patent WO0229103.
 ACCESSION AX411163
 VERSION AX411163.1 GI:21443868
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 1 Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
 Gene expression profiles in liver cancer
 Patent: WO 0229103-A 3810 11-APR-2002;
 JOURNAL GENE LOGIC INC (US)
 FEATURES
 source location/Qualifiers
 1..3654
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="EMBL/Genbank Accession No. X83228"
 BASE COUNT 1052 a 804 c 777 g 1021 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,99e-317 Length: 3654
 Score: 4303.00 Matches: 828
 Percent Similarity: 99.76% Conservative: 2
 Best Local Similarity: 99.52% Mismatches: 2
 Query Match: 99.58% Indels: 0
 DB: 6 Gaps: 0
 US-10-025-380-1081 (1-832) x AX411163 (1-3654)
 Qy 1 MetIleuGlnAlaHisIleuHisSerLeuCysLeuLeuMetLeuTyraIleAlaThrGly 20
 Db 93 ATGATCTTCAAGGCCCATCTTCACTCCCTGTGTCTTATGCTTATTTAGGCACTGCA 152
 Qy 21 TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrglu 40
 Db 153 TATGCGCAAGAGGGGAAGTTTATGACCCCTGAACCCATGACATTTCTATTATTAAGA 212
 Qy 41 GlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAsnProProAlaValThrPhe 60
 Db 213 GCCCAAGAACCGAGTCAATATATTCAGTTTAAGGCCATCTCTGCTGTGACTTTT 272
 Qy 61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluAlaArgGluGlyLeuLeuTy 80
 Db 273 GAACATACTGGGAGACAGCAACATATTTGTATGAACAGGAGGAGGACTTGTGATTAAC 332
 Qy 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaLeuAsp 100

```
|||||
Db 333 AACGAGCCTTGGCAGGAAACAAAGATCTACTACATCTCCAGTTGGAGCCCTGGAC 392
Qy 101 ALaSnglyIleIleValGluGlyProValProIleThrIleGluValAspIleAsn 120
Db 393 GCTAATGAAATTAATAGTGAAGGAGTCCAGTCCATATCCCAATTAAGAAAGGACATCAAC 452
Qy 121 AspAsnArgProThrPheLeuGlnSerIleTyrGluGlySerValArgGlnAsnSerArg 140
Db 453 GACATGACCCACGTTCTCCAGTCAAGTACGAAGGCTCAGTAAGCGAAGACTCTCGC 512
Qy 141 ProGlyLysProPheLeuTyrValAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
Db 513 CCAAGAAAGCCCTCTGTATGTCAATGCCACAGACCTGGATGATCCGGCCACTCCCAAT 572
Qy 161 GluGlnLeuTyrTyrGlnIleValIleGluLeuProMetIleAsnAsnValMetTyrPhe 180
Db 573 GGCCAGCTTATATACCAATGTATCATCCAGCTTCCAGATGACCAATGTCAATGATCTTT 632
Qy 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
Db 633 CAGATCAACAAACAAAGGAGGCATCTCTTATCCGAGAGGAGATCTCAGGAATTGAAT 692
Qy 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlyGlnSer 220
Db 693 CCTGCTAAGATCTCTCTTAAATCTGTATCTCAGTGAAGGACATGGAGGCGCAGAGT 752
Qy 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrLys 240
Db 753 GAGAAATCTTCAAGTATACCAATCTGTGGATATCAATGATGACAGGAATATTGGAAA 812
Qy 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
Db 813 GCACCAAAACCTGTGGAGATGGTGGAAACTCAACTATCTCAACCCCAACAAATCACT 872
Qy 261 GluValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGluLysLeuPro 280
Db 873 CAGGTGGGTGGAATGATCCCGGTGCACAAATATCTTAGTTGACAAAGAGAGCTCCA 932
Qy 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrValThrGlnProLeuAspArg 300
Db 933 AGATTCCTCATTTCAATTGACCAAGAAAGATATTTACGTGACTCAGCCCTTGGACGA 992
Qy 301 GluGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu 320
Db 993 GAGAAAGAGATGCAATATGTTTATGCAAGTTGCAGAAAGATGAGTGGAAAAACACTT 1052
Qy 321 SerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProProThrCys 340
Db 1053 TCATATCCGCTGGAATTCATGTAAAGTTAAAGATATATATATATCATCATCACTACATGT 1112
Qy 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
Db 1113 CCGTCCACAGATTAACCGATTGTAGGTCCAGAGAAATGAAGACTGGGTAAACAGTATCGGG 1172
Qy 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
Db 1173 ACCCTTACTGCAATGACAGGAGTGAAGAAATATCTGCCAACAGTTTTCTTAACTACAGG 1232
Qy 381 IleValGluGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
Db 1233 ATGTGAGAGCAACTCCCAAACTTCCATGATGAGACTTCTTCAATCCAAACTATAGCT 1292
Qy 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyrAsnLeu 420
Db 1293 GGAATGTTCAGTTAGCTAAACAGTCTCTGAAGAGAGATATCTCTCTGATCAACTTA 1352
Qy 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
Db 1353 ACGATAGAGGTGTCTGACAAAGATTTCAAGACCTTTGTTTGTGCAATCAACGTTATT 1412
Qy 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAla 460
|||||
Db 1413 GATATCAATGATCAGACCCCATCTTTGAAAAATTCAGATTATGAAACCTGACTTGTCT 1472
Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGluAlaThrAspAlaAspGluPro 480
Db 1473 GAAAGACAAACAAATGGGTGACCATCTTATCAACATCAGGCGCATGTATGCTGATGAGCCA 1532
Qy 481 PheThrGlySerSerLysIleLeuTyrHisIleIleLysGlyAspSerGluGlyArgLeu 500
Db 1533 TTTACTGGAGCTTCTAAATTTCTGTATCATATATTAAGGAGACAGTGAAGGACCGCTG 1592
Qy 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
Db 1593 GGGGTTGACACAGATCCCATACCAACACCGGATATGTCAATTAATAAAAGCCCTTGAT 1652
Qy 521 PheGluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
Db 1653 TTGAAACGACACTGTTTCCAAATGTTGTTAAAGAGAAATCTTGAGCTCTAGTG 1712
Qy 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
Db 1713 TTTGTGTGAAGTACAAATGCAATGTTCTTTGGCAATTCACGCTTATTTGTGACAGATGTG 1772
Qy 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
Db 1773 AATGAAGCACTCAATTTTCCCAACAGTATTCCAAGCCGAAGTCAGTGAGATGTAAGT 1832
Qy 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTyr 600
Db 1833 ATGAGCACTTAAGTGGCCAAATGATGATGCTGCAGAGATCCAGAAAGTCTGACATTAAGCT 1892
Qy 601 SerLeuArgGlyAspThrArgGlyTyrLeuLysIleAspHisValThrGlyLysIlePhe 620
Db 1893 TCACATAGGGAGACACAAAGAGTTGGCTTAAATGACACAGCTGCTGGTGAATCTTT 1952
Qy 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaThr 640
Db 1953 AGTGTGCTTCATTTGACAGAGAACCGGAAGTCCATATCGGGTACAAGTGGTGGCCACA 2012
Qy 641 GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
Db 2013 GAAGTAGGGGGGCTTCTCTTGAAGCTGTGTGAGAGTTCCACTGATCCCTTAAGATGTG 2072
Qy 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
Db 2073 AATGACAAACCTCCAGGCTAGCCAAAGACTACACGGGCTGTTCTTCTGCCATCCCTC 2132
Qy 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheArgGly 700
Db 2133 AGTGACACTGGAAAGTCTCATTTTTCAGGCTACATGATGATCAGCACTTATTCGGGGT 2192
Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerLys 720
Db 2193 CCCCATTTTAACTTTTCCCTCGCAGTGAAGCTTACAAAGCACTGGGAAGTTTCCAAA 2252
Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaTyr 740
Db 2253 ATCAATGATATCATATCCCACTGCTATCAACAGCACACAGATTTTACAGAGAGGAGCTAT 2312
Qy 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGlnGluIleValSerLeu 760
Db 2313 GTGCTGTGATCCGATCAATGATGGGGGTGGGCCACCTTGGAAAGCAATTTGTTCTTTA 2372
Qy 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
Db 2373 CCAATTACATTTCTGCGTGTGTGGAAGAGTGTTCCTGGCGCAGCAAGTACACAGACT 2432
Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
Db 2433 GGGATACCACTGTGGGCAATGGCAATGGTATATCTGACCAACCTTCTGGATGGAT 2492
Qy 801 IleIleLeuAlaValAlaPheIleArgIleLysLysAspLysGlyLysAspAsnValGlu 820
Db 2493 ATTAATTTAGCAAGTGTGTTATCCGCATTAAGAAAGATTAAGGCAAAAGATATGTTGAA 2552
```

QY 821 Ser1aGln1aSerGluValLeuProLeuArgSer 832
 DB 2553 AGTCTCAAGCATCTGAACTCAACCTCTGAGAAC 2588

RESULT 11
 HSRNALICA

LOCUS HSRNALICA 3654 bp mRNA linear PRI 01-JUN-1995
 DEFINITION H.sapiens mRNA for LI-cadherin.
 ACCESSION X83228
 VERSION X83228.1 GI:854174
 KEYWORDS LI-cadherin.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS 1 (bases 1 to 3654)
 Boettlinger, A., Krefte, B., Fieger, C., Dlouhy, B., Berndorff, D.,
 Goessens, R., and Tauber, R.
 Molecular cloning of human LI-cadherin: evidence for a novel type of
 cadherin within the cadherin superfamily

TITLE Unpublished
 2 (bases 1 to 3654)
 Boettlinger, A.M.
 Direct Submission
 Submitted (05-DEC-1994) A.M. Boettlinger, Institut fuer Klinische
 Chemie & Bioch., Universitaetsklinikum Rudolf Virchow, Freie
 Universitaet Berlin, Spandauer Damm 130, 14050 Berlin, FRG

COMMENT
 FEATURES
 source
 1. 3654
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="uni-ZAP TM vector"
 93. 2591
 /function="cell adhesion molecule"
 /codon_start=1
 /product="LI-cadherin"
 /protein_id="CAA8231.1"
 /db_xref="GI:854175"
 /db_xref="SPTRMBL:Q15336"
 /translation="MILQAHLSLCLMLYLATYVGQEGKFSGLKPMPTSIYEGQEP
 SOIFQKAMPATVPELTGETDNIPIYIEBGLLYRALDREHRSHTNVALADN
 GIIYEGVPITIKDINDRPFLOSIGSVGRONRPPKPIYVNAATLDDAPFN
 GQLYQIVIQPLMNNMYROIINKTGALISITREGSELNPAKPSINVIYSKDMG
 QSENSFSDTTSVDLIITENIKWAKPVEWENSTDPHIIKTQVRMDPAQSLVDK
 EKLPRFSDIDOBGDIYVQPLDRBEDAVYFAVADEGKPLSYLEIHKVKDIN
 DNPTQSPVTVFEVQENRIGNSIGTLTAHDRENTANSFLNRYIVEQTPKLMG
 LFIQTAGMQLAKGSLKQDTPQYNLTIEVSKDPTKICFQVINDINDPTRE
 KSDYGNITLADNTNIGSTLTITQATDADPEPTGSKLHYTIKDSRGLGVNDPPT
 NTGYVITIKRPLDEFIAVSNIVFAENRBPPLVGVKTNASSFAFTLIVDNEAPF
 SQHVFQAKVEDVAIGTKVGNVAKDEBGLDISYLSGDKTRGMLKIDHVGELTSVA
 LDRAGSPRYOVAVTEVGSLSVSEFHLIMDVANDNPRKLADYTGCFEFCFLSA
 PGSILFATDDOHLFRGPHETFSLSGSQNDWEVSKINGTHARLSRHEFEREY
 VVLRINDGGRPLEGIVSLPVTCSYEGSCFRPAGHQGIPVMAVAGLLTLTLV
 IGLTAVFIRIKDKGNVESQASEVKEPLRS"

BASE COUNT 1052 a 804 c 777 g 1021 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.99e-317 Length: 3654
 Score: 4303.00 Matches: 828
 Percent Similarity: 99.76% Conservative: 2
 Best Local Similarity: 99.52% Mismatches: 2
 Query Match: 99.58% Indels: 0
 DB: 9 Gaps: 0

US-10-025-380-1081 (1-832) x HSRNALICA (1-3654)

QY 1 Met1leuGln1aH1sleuH1sSerLeuCysleuLeuMetLeuYrLeuAlaThrGly 20
 DB 93 ATGATCTTCAGGCCATCTTCACCTCGTGTCTTTATTGAGCTTATTGGCAACATGGA 152

QY 21 TYRGLYGLNGLUGLYLYSPHESERGLYPROLEULYSPROMETHRPHESERIETRYGLN 40
 DB 153 TATGCCCAAGAGGGGAAAGTTAGTGAGACCCCTGAAACCATGACATTTCTATTATTGAA 212

QY 41 GLYGLNGLUPROSERGLNILEPHEGLINPHELYALASNPROFOLAVALTIRPHE 60
 DB 213 GGCCAGAAACCGGATCAAAATTATATTCAGATTAAAGCCAAATCCTCTGAGACTTTT 272

QY 61 GILUETHRGLIGLUHRPASPANILEPHEVALILEGLUARGGLUGLYLEULEUYTRYX 80
 DB 273 GAACACTACGGGAGACAGCAACATATTGTGATGAACGGAGGACTTCTGTATTAC 332

QY 81 ASNARGALALEUASPARGLUHRARGSERTHRHAASNLGULVALALALEUASP 100
 DB 333 AACAGAGCCTTGACAGGGAACAAATCTACTCAAACTCCAGGTGAGCCCTGGAC 392

QY 101 ALAASNGLYILELEVALIGLUGLYPROVALPROILETHRILEGLUVALYASPILEA 120
 DB 393 GCTAATGGAATTATAGTGGAGGCTCCAGTCCCTATCACAATAAAGTGAAAGCATCAG 452

QY 121 ASPASNARGPROTHRPHLEUGLINSERLYTRYGLUGLYSERVALARGGLNANSEARG 140
 DB 453 GACAAATGACCCACGTTTCTCCAGTCAAGTACGAAGCTCACTAAGCGAAGCTCCG 512

QY 141 PROGLYLYSPROPHLEUYTRYVALAASNALATHRASPLEUASPARGPROALATHRPROAS 160
 DB 513 CCAAGAAAGCCCTCTTGATATGATGATGACAGACATGATGATGATGATGATGATGAT 572

QY 161 GLYGLNGLUYTRYGLNILEVALILEGLINLEUPROMETILLEASNPANVALMETTRYPHE 180
 DB 573 GGCCAGCTTATATCAAGATTCATCAGCTCCATGATCAATGATCAATGATCAATGAT 632

QY 181 GNLILASNPANLYSHTRGYALILESERLEUTHRARGGLUGLYSERGLNGLULASNP 200
 DB 633 CAGATCAACAAACAAAGGAGCATCTCTTACCCGAGAGGAGATTCAGGAATTGAAT 652

QY 201 PROALALYASNPROSETRYRPAENLEUVALILESERVALYASNPMEGLYGLINSER 220
 DB 693 CTGCTNAGAACTCTCTGATATGATGATGATGATGATGATGATGATGATGATGATGAT 752

QY 221 GLUASNPESRPHESERAPTHRTHRSERVALASPILELEVALTHRGLUASNILETRPLY 240
 DB 753 GAAATTCCTTCAGTATACCAATCTGATGATGATGATGATGATGATGATGATGATGAT 812

QY 241 ALAPROLYSERVALIGLUETVALIGUASNPESRTHASNPROMISPROILELYSILETH 260
 DB 813 GCACCAAAACCTGTGAGAGTGGAAACCTCAACGATCTCAACCCATCAAAATCACT 872

QY 261 GILVALARGTRPASPANPPROGLYALAGINTYRSELEUVALASPLYSGULYSPLEU 280
 DB 873 CAGGTGGGTGGAATGATCCCGGTGACAAATATCTCTTATGATCAAAAGAAAGCTGCCA 932

QY 281 ARGPHERPHESERILEAPGLINGLUGLYASPILETRYVALTHRGLNPROLEUASPARG 300
 DB 933 AGATTCCCATTTCAATTGACCGAAGAGATATTACCTGACTGACCCCTTGACCGGA 992

QY 301 GLUGLUYASNPALATRYVALPHETRYVALALALYASNPGLUYTRYGLYSPROLEU 320
 DB 993 GAAAGAAAGATCATATGTTTATTATGACGTTGCAAGATGATGATGATGATGATGAT 1052

QY 321 SETRYRPROLEUGLUILEHISVALYVALYASPILEASNPASPANPROFOLTHCYS 340
 DB 1053 TCATATCCGCTGGAATTCATGTAAGATTAAAGATTATTAATGATCAACCATCATGT 1112

QY 341 PROSERPROVALTHRVALPHEGLUVALGLINGLUASNGLUARGLEUGLYANSERILEG 360
 DB 1113 CCGTCAACAGTAACCGATATTGAGTCCAGAGAAATGAACAGCTGGGTAAACAGTATCGG 1172

QY 361 THRLEUTHRALHISAPARGASPGLUGLUASNPTRHALASNPSPHELEUANTYTRYARG 380
 DB 1173 ACCCTTACTCACAATGACAGGATGAAGAAATATCTGCCAAGTTTCTTAACTACAGG 1232

QY 381 ILEVALIGLUGLINTHRPROLYSPLEUPROMETASPGLYLEUPHELEULEGINTHRTRYALA 400

Db 1233 ATTGTGAGCAAACTCCCAACTTCCTCCATGATGAGCTCTCCATATCCAAACCTATGCT 1292
 Qy 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnIleThrAspLeu 420
 Db 1293 GGAATGTTACAGTTAGCTTAAACAGTCTCTGAGAGAGCAAGTACTCTCTGATACACTTA 1352
 Qy 421 ThrIleGluValSerAspLysAspPheLysThrLeuGlnValGlnIleLeuValIle 440
 Db 1353 ACAGTACAGAGGTGTGACAAAGATTTCAGACCTTTGTTGTCGAAATCAAGTTATT 1412
 Qy 441 AspIleAsnAspGlnIleProIlePheGlnLysSerAspTyrGlyAsnLeuThrLeuAla 460
 Db 1413 GATATCAATGATACAGACCCCATCTTGAATAAATCAATTAAGAAACCTGACCTTGCT 1472
 Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnIleAlaThrAspAlaAspGluPro 480
 Db 1473 GAAAGCAAAACATGGGTGTCACATCTTAACATCCAGGCACTGATGCTGATGAGCCCA 1532
 Qy 481 PheThrGlySerSerLysIleLeuThrIleLysIleIleLysGlyAspSerGlnGlyArgLeu 500
 Db 1533 TTTTACTGGAGATTCTAAATTTCTGTATCATATCTTAAAGGAGACAGTGAAGGACGCTG 1592
 Qy 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
 Db 1593 GGGGTGACACAGATCCCATACCAACACCGGATATGTCATTAATTAAGCCCTTGAT 1652
 Qy 521 PheGlnThrAlaAlaValSerAsnIleValPheValAlaGluAsnProGluProLeuVal 540
 Db 1653 TTTGAAACACAGCTGTTTCCAAATGTTGTTCCAAACAGAAATCTTGAGCCCTTAGTG 1712
 Qy 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValIleThrAspVal 560
 Db 1713 TTTGGTGTGAGATGACATGACAGTTCTTTGGCCAGTTCCGCTTATTTGACAGAGTGT 1772
 Qy 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 Db 1773 AATGAAGCACCTCAATTTTCCCAACAGTATCCAAACGAAAGTCAAGTGAAGTGTAGCT 1832
 Qy 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluLysLeuAspIleSerTyr 600
 Db 1833 ATAGGCACTAAGGGGCAATGATGCTCCAGAGATCCAAAGGTCTGAGCATTAAGCTAT 1892
 Qy 601 SerLeuArgLysAspThrArgGlyTyrLeuLysIleAspHisValThrGlyLysIlePhe 620
 Db 1893 TCCTGAGGGGAGACACAAAGGTGCTTAAATTCACACGCTGAGTGTGAGATTTT 1952
 Qy 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaIleThr 640
 Db 1953 AGTGTGCTCCATTTGACAGAAAGCCGAAAGTCCATATCGGGTACAAAGTGTGGCCACA 2012
 Qy 641 GluValGlyLysSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 Db 2013 GAAGTGGGGGGGCTCTCTTCTGAGCTGTGTCAGAGTTCACCTGATCCTTATGATGTG 2072
 Qy 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 2073 AATGACAAACCTCCAGGCTAGCCAGAGACTTACACGGGCTTGTCTTCTCCATCCCTC 2132
 Qy 681 SerAlaProGlySerLeuIlePheGlnAlaThrAspAspAspGlnHisLeuPheArgGly 700
 Db 2133 AGTGCACCTGAGAGTCTCATTTTCGAGGCTACGATGATGATCAGACCTTATTTCCGGGT 2192
 Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerLys 720
 Db 2193 CCCCATTTTACATTTTCCCTGCGCAGTGAAGGCTTCAAAACACACCTGGGAAAGTTTCCAA 2252
 Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGlnGluAlaTyr 740
 Db 2253 ATCAATGATCTATGCTGCTGCTACAGGACACAGAGTTTGAAGAGAGGAGTAT 2312
 Qy 741 ValValLeuIleArgIleAsnAspGlyArgProProLeuGlnGlyIleValSerLeu 760

Db 2313 GTGCTTGTATCCGATCAATGATGGGGTTCGGCCACCTTGGAGGCAATGTTCTTTA 2372
 Qy 761 ProValThrPheCysSerCysValGluIleSerCysPheAspProAlaGlyHisIleThr 780
 Db 2373 CCAATTATCATTTCTGCGAGTTGTGTGAAGAGTTGTTCCGGCAGAGGCTCCAGACT 2432
 Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2433 GGGATACCCACCTGGGCAATGGCAGTTGGTATCTGTCACACCTCTGCTGATTTGT 2492
 Qy 801 IleIleLeuAlaValAlaPheIleArgIleLysLysAspLysGlyLysAspAsnValGlu 820
 Db 2493 ATATTTTACAGCTTGTGTTATCCCATTAAGAGATTAAGAGCAAAAGATTAATGTGA 2552
 Qy 821 SerAlaGluAlaSerGluValLysProLeuArgSer 832
 Db 2553 AGTGTCAAGCATCTGAAGTCAAACTCTGAGAGC 2588
 RESULT 12
 AF177669 3351 bp mRNA linear ROD 05-JUN-2001
 LOCUS AF177669
 DEFINITION Mus musculus liver-intestine cadherin mRNA, complete cds.
 ACCESSION AF177669
 VERSION AF177669.1 GI:5762489
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS Angres,B., Kim,L., Jung,R., Gessner,R. and Tauber,R.
 TITLE Li-cadherin gene expression during mouse intestinal development
 JOURNAL Dev. Dyn. 221 (2), 182-193 (2001)
 MEDLINE 21270055
 PUBMED 11376485
 REFERENCE
 2 (bases 1 to 3351)
 AUTHORS Angres,B., Kim,L. and Tauber,R.
 TITLE Direct Submission
 JOURNAL Submitted (14-AUG-1999) Institut fuer Klinische Chemie und Biochemie, Virchow-Klinikum der Humboldt-Universitaet zu Berlin, Augustenburger Platz 1, Berlin 13353, Germany
 FEATURES
 source
 1. .3351
 /organism="Mus musculus"
 /strain="Swiss Webster/NIH"
 /db_xref="taxon:10090"
 /cissue_type="intestine"
 /dev_stage="17 day embryo"
 44. .2527
 /note="Li-cadherin, cell-cell adhesion molecule"
 /codon_start=1
 /product="liver-intestine cadherin"
 /protein_id="AAD51125.1"
 /db_xref="GI:5762490"
 /translation="MWSAQLHFLCLITLTLTCGYSEBEGKFSGLKMPMSIFEGEBS
 OYRQKMPAPVPELCEGDCIKREMDGLHYTRADRETRVHHQLALDSHG
 ALVDGVPRTTIVKIDNRPFLLSKTEGSTRQNSRPEKPMYNATLDLPDPSG
 QLEQVIVQLPQIDNVMPFQIDSKGALSLTEGSEQLDPVNPSPYNLVSXWDMGQ
 SENSPSDTIVDISIRENIMKAPBEVETIRENSTDPHPIKITOVQWNPDAQSLVKE
 KLSPPFSIDCGNLYVQALDREKNSHVPATADENGKPLAYLEIYVAVIDND
 NPCTGSPVTFEVOENELGNSIGIFEAHMDANNINSLIKYLVDQTPVPSGL
 FLIGEGEVQVLSKQSLKQDSPOYNLSIEVSDNDPKTCYQVAVIDNDQIPFET
 SNYKTSLESDTALISITLITATADDEFSTSSKILKIVAGDVBGLVETDPTTN
 AGYVKKIKPLDETPVPSIVQANPELVGILVYNASSFASFEILVAVDVEVPF
 QRIFOANVSQAAVAGSVGNVARDPEGLTVYSLSKNNRGMKILDSVGEI.FSAPL
 DRETESYRVQVAAVEGSSLSSTADFLVLDVNDNPPRLAKDYTGIF.PCHSAP
 GSLIFEVDDQDQSLRRPEKTPALRGEGGQMEVSKINGTARLSTHRTREBOVY
 IPIRINDGGQPMEGTVFLPYTFRCQGVGSCRPAGRDGIPTYGMAAGIILLTFLVI
 GIILVVFIRMKDVENPQSPENKPLAS"
 BASE COUNT 966 a 748 c 746 g 891 t
 ORIGIN
 Alignment Scores:

Pred. No.: 1,33e-252 Length: 3351
 Score: 3450.00 Matches: 658
 Percent Similarity: 88.21% Conservative: 75
 Best Local Similarity: 79.18% Mismatches: 94
 Query Match: 79.84% Indels: 4
 DB: 10 Gaps: 1

US-10-025-380-1081 (1-832) x AF177669 (1-3351)

QY 2 ILEuGlnAlaHisLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGlyTyr 21
 DB 44 ATGGTGTGCGCCAGCTTCACTTCTGTCTCTCTCACTTATTTAGCTTGAGCTGATAT 103
 QY 22 GYGLnglGlnGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGlnGly 41
 DB 104 GGGGAAAGAGGAAAGTTCCAGCGGTCCCTGAAAGCCAGAGCATTTCTCATTTTGAAGGC 163
 QY 42 GlnGluProSerGlnIleIlePheGlnPheLysAlaAspProAlaValThrPheGlu 61
 DB 164 CAAAGAACCGAGTCAAGTCATATTCATTTTAAAGCAACCTCCAGCTGAGCTTTCGAG 223
 QY 62 LeuThrGlyGlnThrAspAsnIlePheValIleGluArgGlnGlyLeuLeuTyrTyrAsn 81
 DB 224 CTAAACGGAGAGCAGATGATATTTAAAGATGAAAGATGAGATCTTGATCACACA 283
 QY 82 ArgAlaLeuAspArgGlnThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAspAla 101
 DB 284 AGAGCCCTGGACAGAGAAACAAAGACGGTTCATCATCTGACGCTTGAGCTTGATTC 343
 QY 102 AsnGlyIleIleValGlnGlyProValProIleThrIleGluValLysAspIleAsnAsp 121
 DB 344 CATGAGCTATAGTGAACGGTCCAGTCCCATCCATAGAGTCMAAGCATCAATGAC 403
 QY 122 AsnArgProThrPheLeuGlnSerLysTyrGlnLysSerValArgGlnAsnSerArgPro 141
 DB 404 AACGACCCAGCTTCTCCAGTCAAAATATGAAAGCTCAGAGGCGCAACTCCGCCCA 463
 QY 142 GlyLysProPheLeuTyrValAsnAlaThrAspLeuAspProAlaThrProAsnGly 161
 DB 464 GGAAGGCTTTCAGATGATGTCATGCTAAGTCAAGATCTGATGATCCGCGTACCTCCAAATGGC 523
 QY 162 GlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPheGln 181
 DB 524 CAGCTTTTATTCAAATTTGATCATCTCACTTCCCAATCAACGATTTATGATCTTTCA 583
 QY 182 IleAsnAsnLysThrGlyAlaIleSerLeuThrArgGlnLysSerGlnGlnLeuAsnPro 201
 DB 584 ATCAGACAGCAAAACAGGGGCAATATCGCTTACCCAGAGAGATCCCGAATTTGGATCCA 643
 QY 202 AlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnSerGlu 221
 DB 644 GTGAGAGATCTCTCTCTCAACCTGTGTCTCGGTGAAGAGCATGGGGGCGCAAGTGA 703
 QY 222 AsnSerPheSerAspThrThrSerValAspIleIleValThrGlnAsnIleTyrLysAla 241
 DB 704 AATTCCTTCACTGATCTACATATGATGATATTTCTATTAAGAGAAATATCTGAAAGCA 763
 QY 242 ProLysProValGlnMetValGlnAsnSerThrAspProHisProIleLysIleThrGln 261
 DB 764 CCAAGACCCGCGAGATTAGAGAAACCTGACTGATCTCAACCCATCAAAATCACTCAG 823
 QY 262 ValArgThrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGlnLysLeuProArg 281
 DB 824 GTGAGATGAGATGAGCCAGGGGCGCAGATTCCTTAAGTCAACAGAGAAACCTGCGCG 883
 QY 282 PheProPheSerIleAspGlnGlnGlyAspIleTyrValThrGlnProLeuAspArgGlu 301
 DB 884 TTCCCATTTCTGATCGACCAAGAGAAATATTTATGATCTCAGGCTTGACCGGAG 943
 QY 302 GlnLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeuSer 321
 DB 944 GAAAGAACTCACTGTTTCTTTGCACTGCGAAGGATGAGAAATGAGAAACCACTTGCA 1003

QY 322 TyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProThrCysPro 341
 DB 1004 TATCCACTGGAATTTATGAAAGTATTTAGCATTTATGACAAACCACTGATCTG 1063
 QY 342 SerProValThrValPheGluValGlnGluAsnGluArgLysGlyAsnSerIleGlyTyr 361
 DB 1064 TCTCCAGTACGTATTTGAGTCCAGGAGAAATGAAACATTTGGTAAACGATCGGATC 1123
 QY 362 LeuThrAlaHisAspArgAspGlnGluAsnThrAlaAsnSerPheLeuAsnTyrArgIle 381
 DB 1124 TTGAGGCCCAATGATGAGACGAACTTAACATCAACATGATTTTGAATTAACAGCTT 1183
 QY 382 ValGlnGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAlaGly 401
 DB 1184 GTAGACCAAAACCCCAAGTTCCCTCAGATGAGCTTTCTCATGCTGTAATGAGGGA 1243
 QY 402 MetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyrAsnLeuThr 421
 DB 1244 AAGGTTCACTTAAGTAAACAGTCTTGAAAGAGCAAGACAGCTCTCAGTACAACTTAAGT 1303
 QY 422 IleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIleAsp 441
 DB 1304 ATGAGAGTGTGACCTGAATTAATTCAGACTCTGTATATTTCAAGTCAACGTTATGAT 1363
 QY 442 IleAsnAspGlnIleProIlePheGlnLysSerAspTyrGlyAsnLeuThrLeuAlaGlu 461
 DB 1364 ATCAATGATCAAGTTCCATCTTTGAACATCAATTAAGAAAGCAAGCTGCTGTA 1423
 QY 462 AspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluProPhe 481
 DB 1424 GACACCGCATCGAGTCCACATCTTAATCATCAAGTCAAGATCTGATGAGCCCTT 1483
 QY 482 ThrGlySerSerTyrIleLeuTyrHisIleIleLeuGlyAspSerGlnGlyArgLeuGly 501
 DB 1484 ACAGGAGGCTTAATAATCTGTAACAGATTTGACAGAGACACGAGGAGAACTTAGAA 1543
 QY 502 ValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysProLeuAspPhe 521
 DB 1544 GTTGTCAAGATTCACAGCAACATGCAAGATATGTCMAAAGCTCTTGATTTT 1603
 QY 522 GlnThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuValPhe 541
 DB 1604 GAAACTCAACCACTTGACAGCATTTGTTCCAAAGCGGAAATCTGAGCCACTGGGAAA 1663
 QY 542 GlyValLysThrAsnLysSerPheAlaLysPheThrLeuIleValThrAspValAsn 561
 DB 1664 GGCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1723
 QY 562 GlnAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAlaIle 581
 DB 1724 GAAAGTGTGTGTTCCCGCCAGCGAATATTCMAAGCAAGCTCAGCGAGATGCTGTGTG 1783
 QY 582 GlyThrLysValGlyAsnValThrAlaLysAspProGlnGlyLeuAspIleSerTyrSer 601
 DB 1784 GCGAGTAGAGTGGCAACGTGATGCGCAGGATCTCGAAGCGCTGACTGAGATTAATCA 1843
 QY 602 LeuArgGlyAspThrArgGlyTyrPheLysIleAspHisValThrGlyGlnIlePheSer 621
 DB 1844 CTGAAGGCAATATGAGAGGCTGCTTAATATGACTCTCTCACTGCTGAGATTTAGT 1903
 QY 622 ValAlaProLeuAspArgGlnAlaGlySerProTyrArgValGlnValAlaIleThrGlu 641
 DB 1904 GCTGCTCCGTGTGATGAGGAAACAGAAAGTGTATTCGATTAACAAGTGGGCGACATGAA 1963
 QY 642 ValGlyLysSerSerLeuSerSerValSerGlnPheHisLeuIleLeuMetAspValAsn 661
 DB 1964 GTAGTGGTGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2023
 QY 662 AspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeuSer 681
 DB 2024 GACAAACCCCTGCTGATGAGACTTAACAGGCTGTTGTTCTTCTGCAATCCCTCAGT 2083
 QY 682 AlaProGlySerLeuIlePheGlnAlaThrAspAspAspGlnHisLeuPheArgGlyPro 701


```

|||||
Db      2084  GCGCCCTGGAGCCCTCATCTTTGAGGTCACATGACCATGACGACATGCTTCGAGAGGCC 2143
|||
Qy      702  HspHetherPheSerLeuGlySerGlySerLeuInaenAptPgluValSerLeuGly 721
|||
Db      2144  AAGTTTACATTCGCGCTTGGAGGAAAGGCTTCAAAAGTGACTGGGAAAGTTTCAAAATC 2203
|||
Qy      722  AsnGlyThrHisAlaAglLeuSerThrArgHisThrAspPheGluGluArgAlaTyrVal 741
|||
Db      2204  AATGGTACACATGTGTAACTCTCCACGACGACACACGCTTTGAGGAACAAGTTTAAAC 2263
|||
Qy      742  ValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSerLeuPro 761
|||
Db      2264  ATTCCAAATCCGCAATGATGAGGGGCGACGCCATGAGAGGAGGACTGCTTCTTACCA 2323
|||
Qy      762  ValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThrGly 781
|||
Db      2324  GTTACTTCTTGCACAGTGTGTGAAAGAAAGCTGTTCCGCGACGACGACGACGATGGG 2383
|||
Qy      782  IleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGlyIle 801
|||
Db      2384  ATACCCACTGTGGGCAATGAGGAGTGTGTACTTCTGACACATTTCTGCTCATTTGATTA 2443
|||
Qy      802  IleLeuAlaValValPheIleArgIleLeuLeuAspLeuGlySerValAspAsnValGlySer 821
|||
Db      2444  ATTTTACAGTGTCTTCATCCGACGACGAGAAAGATATAA-----GTTGAAAT 2491
|||
Qy      822  AlaGlnAlaSerGluValSerProLeuArgSer 832
|||
Db      2492  CCGAGATCCCTTGAAGATTAAGCTCTTGAAGAGC 2524
|||
RESULT 13
Db      2492  CCGAGATCCCTTGAAGATTAAGCTCTTGAAGAGC 2524
|||
LOCUS   D87912          3414 bp      mRNA      linear      ROD 06-OCT-2000
DEFINITION Mus musculus mRNA for BHLH-cadherin, complete cds.
ACCESSION D87912.1 GI:9392287
VERSION   D87912.1
KEYWORDS  Mus musculus (strain: BALB/c) foetal liver Prob cell cell_line:3889
SOURCE    CDNA to mRNA.
ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3414)
AUTHORS   Ohnishi, K., Shimizu, T., Karasuyama, H. and Melchers, F.
TITLE      The identification of a nonclassical cadherin expressed during B
            cell development and its interaction with surrogate light chain
            J. Biol. Chem. 275 (40), 31134-31144 (2000)
JOURNAL   20469471
MEDLINE   2
REFERENCE 2 (bases 1 to 3414)
AUTHORS   Ohnishi, K.
TITLE      Direct Submission
            Submitted (18-SEP-1996) Kazuo Ohnishi, National Institute of
            Health, Department of Immunology, 1-23-1, Toyama, Shinku-ku,
            Tokyo 162, Japan (E-mail: ohnishi.k@nih.go.jp,
            Tel: 03-5285-1111 (ex. 2133), Fax: 03-5285-1150)
FEATURES
     source
         1..3414
             /organism="Mus musculus"
             /strain="BALB/c"
             /db_xref="taxon:10090"
             /cell_line="3889"
             /cell_type="Prob cell"
             /tissue_type="foetal liver"
             /codon_start=1
             /product="BHLH-cadherin"
             /protein_id="BAB03264.1"
             /db_xref="GI:9392288"
             /translation="MYSQAQHFICLTLVITCGYGEKGRSGPLKWTSTISIFRGQPS
             QVIFQRINPVAATFELTGETDGIKFKIKDGLYHTRALDRFRAVHHQLALDSHG
             ALVDGPVITIEVKDINDNRPFLQSKYESVRQNSRPGKPFMYVAVATLDDPATNG
             QLFYQIVQLFQINDVMYFQIDSKTASLSTPGSQELDPVKPNSYLVVSVKMDGSG

```

```

SENSFSDTIVDISIRENIMWKAPEVEIRENSTDPPIKITOVQWMDPGAQVLSVKE
KLSPPSPSIDOENIYVTOALDEEKGSHVFPATADENKPLAYPEIIVKYKIND
NPTCLSPVTFVQENEPNIGSIIFEAHDMENANNISLKYKLVDPKRPNSGL
FLIGVEBKQVLSKSLKODSQRNLSIEVSVDKRTCTCYTOVNVIDIDQPIRET
SNYGSKTLSDEDTALIGSTLITIOATDDEPTGSSKILYKIVQDTEGRLEVDPTTN
AGVYKIKLSPDEPOTVSSIVFOAENPEPLVKIEVNASPEAFELIVDNDVVPFP
QRIFOANVSEDAVGSRVGNVTADEEGELVSYSLGNMNGMKIDSVGEISAPL
DRETSVYRQVAVTGVSSLSSTADPHLVLTVDVNDNPRKLADYTGEPFPLSP
GSLIFVETDDQOGLARPKPTFALREGIOSEVBSKINGTNTHTSTHTRFEEQVYN
IPRINDGGQPFMEGVTFPLVTECCQVBBSCFRPAQRDSIPTVGNVAILTLTFVLV
GIILAVFTIMRDXENPQSPENKPLRS"
BASE COUNT      967 a      766 c      774 g      907 t
ORIGIN
Alignment Scores:
Pred. No.:      1,36e-252      Length:      3414
Score:          3450.00      Matches:      658
Percent Similarity: 88.21%      Conservative: 75
Best Local Similarity: 79.18%      Mismatches: 94
Query Match:      79.84%      Indels:      4
DB:              10      Gaps:      1
US-10-025-380-1081 (1-832) x D87912 (1-3414)
Qy      2  IleLeuGlnAlaHisLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGlyTyr 21
|||||
Db      93  ATGGTGTGCGCCACGCTTCACTTCCCTGTGTCTTCTACACTTATTTGACCTGTGATAT 152
|||||
Qy      22  GlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGluGly 41
|||||
Db      153  GCGGAAGAAGGGAAGTTCAAGCGGTCCCTGAAAGCCCATGACATTTCCATTTTGAAGGC 212
|||||
Qy      42  GlnGluProSerGlnIleIlePheGlnPheValAlaAsnProProAlaValThrPheGln 61
|||||
Db      213  CAAGAACCAGTCAAGTCAATTCAGTTTAAACCAACCTTCACCTGTGACTTTCGAG 272
|||||
Qy      62  LeuThrGlyGluThrAspAsnIlePheValIleGluArgGluGlyLeuLeuTyrTyrAsn 81
|||||
Db      273  CTAAACGGAGACACATGTGTATTTAGATAGAAAGATGAGACTTCTGTATCACACA 332
|||||
Qy      82  ArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAspAla 101
|||||
Db      333  AGAGCCCTGGACAGAGAAACAAGAGCGTTCAATCATCTGACGCTTGACGCTTGATTC 392
|||||
Qy      102  AsnGlyIleIleValGluGlyProValProIleThrIleGluValLysAspIleAsnAsp 121
|||||
Db      393  CATGGAGCTATAGTGAAGCGTCCAGTCCCATCAACATGAAATGACATCAATGAC 452
|||||
Qy      122  AsnArgProThrPheLeuGlnSerIleTyrGluGlySerValArgGlnAsnSerArgPro 141
|||||
Db      453  AACCGACCAACGCTTCTCCAGTCAAAATGAAAGCTCAGTGAAGCAAACTCTCCGCCA 512
|||||
Qy      142  GlyLysProPheLeuTyrValAsnAlaThrAspLeuAspAspProAlaThrProAsnGly 161
|||||
Db      513  GGAAGACCTTTCATGATGACATGCTCAAGATCGATGATCGGCTACTCCCAATGGC 572
|||||
Qy      162  GlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPheGln 181
|||||
Db      573  CAGCTTTTATTCAAATGTCTATCCCACTTCCCAATCAACATCAATGATTTATGTACTTCA 632
|||||
Qy      182  IleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsnPro 201
|||||
Db      633  ATGACAGCAAAACAGGGGCAATATGCTTACCCCAAGAGATCCAGAAATTTGATTCAC 652
|||||
Qy      202  AlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnGlnSerGln 221
|||||
Db      693  GTGAAGAATTCCTTCTCAACACCTGGTGTCTCGTGAAGAGACATGGGGGCGCAAGTGA 752
|||||
Qy      222  AsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrPlyAla 241
|||||
Db      753  AATTCCTTCAAGCATCTCATATGATGATATTTCTATTAGAGAAATATCTCGAAAGCA 812
|||||
Qy      242  ProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThrGln 261

```



```

Db      1828 GTTGTACAGATCCCAATGACCAATACAGATACGATCAGATCAGAAAGCCTTGAATTT 1887
Qy      522  GluThrAlaAlaValSerAsnIleValPheValAlaGluAsnProGluProLeuValPhe 541
Db      1888 GAAACCGAACAGTACTACATGCTGTCAAGCAAGAAACCTGAGCGCGTGTGAAT 1947
Qy      542  GlyValIleValIleValSerSerPheAlaIlePheThrLeuIleValIleThrAspValAsn 561
Db      1948 GGCATACAGTACATAGTCTGCTTCTTTCCTCCCTCAGCTGACCGTACAGAGATGAAT 2007
Qy      562  GluAlaProGlnPheSerGlnIleValPheGlnAlaIleValSerGluAspValAlaIle 581
Db      2008 GAAAGTCTGTATTTTCCCGACCAATATTCACAGCCCAAGCTTACCGAGATCTGCTATA 2067
Qy      582  GlyThrIleValGlyAsnValIleThrAlaIleAspProGluGlyLeuAspIleSerTyrSer 601
Db      2068 GGCATTAAGTGGGCACTGTGATCGCAGGATCCGAAAGGCCCTGACCTGAGATTATTC 2127
Qy      602  LeuArgGlyAspThrArgGlyTyrPheLeuIleAspIleValIleThrGlyGluIlePheSer 621
Db      2128 CTGAAACACATATAGAGAGGTGGCTTAAATGACCTCCGACCGGCGACATATTTAGC 2187
Qy      622  ValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaIleThrGlu 641
Db      2188 ACTGCTCCGCTGACAGGAAACAGAAAGTGTATGAGTACAAAGTGTGGCCACTGAA 2247
Qy      642  ValGlyIleSerSerIleSerSerValSerGluPheIleLeuIleLeuMetAspValAsn 661
Db      2248 GTAGTGGGTCTCTGAGATTCTACAGCATATTTCCACCTGGCTCATGAGATGAAT 2307
Qy      662  AspAsnProProAlaGluAlaIleAspTyrThrGlyLeuPhePheCysIleProLeuSer 681
Db      2308 GACAAACCCCTCCCTGAGCTTACAGACTACACAGCTTGTCTTCCTCCATCCCTTAGT 2367
Qy      682  AlaProGlySerLeuIlePheGlnAlaIleThrAspAspGlnIleIleLeuPheArgGlyPro 701
Db      2368 GCCCTCGGAGCCCTCATCTTGAAGCCATGATATGACCAAGTCAAGTCCGAGGCC 2427
Qy      702  HisPheThrPheSerIleGlySerGlySerLeuGlnAsnAspTyrGluValSerIle 721
Db      2428 AAGTTTACGTTTGCCCTTGGCAGAGAAAGCTTACAAAGTGCAGGAAGTGTCCAAATC 2487
Qy      722  AsnGlyThrHisAlaIleGluSerThrArgHisThrAspPheGluGluAlaIleVal 741
Db      2488 AATGACACATGTGTAGCTCTCCACCAAGCACACGCTTGTGAGAACAGTTATGAC 2547
Qy      742  ValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSerLeuPro 761
Db      2548 ATCCCAATCCCTCATCATGATGAGGGGAGCCACCCATGAGAGCATTTCTCTTATCA 2607
Qy      762  ValThrPheCysSerCysValGlyIleGlySerCysPheArgProAlaGlyHisGlnThrGly 781
Db      2608 GTTACTTTCTGCGCAGTGTGAGAGGAAGTTGTTCCGCGCAGCAGGCAACAGAGTTGGG 2667
Qy      782  IleProThrValIleGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGlyIle 801
Db      2668 ATACCCACCGTGGCAGTGGCAGTGTGTACTTCTGACACATTTCTGTCATTGTGATA 2727
Qy      802  IleLeuAlaValIlePheIleArgIleValIleValAspIleValAspAsnValIleGlyIle 821
Db      2728 ATTTTACCGTGTGATTTCTATCCGTATGAGAAAGATATA-----GTTGAAGAT 2775
Qy      822  AlaGlnAlaSerGluValIleProLeuArgSer 832
Db      2776 CCACAGTCCCTGAGAAATTAAGCTCTGAGAGC 2808

```

```

RESULT 15
RATDPRPT      1733 bp      mRNA      linear      ROD 14-DEC-1995
LOCUS          Rattus norvegicus proton-driven peptide transporter mRNA, 3' end of
DEFINITION
cde.
ACCESSION     L46874

```

```

VERSION      L46874.1 GI:1082042
KEYWORDS     proton-driven peptide transporter.
SOURCE       Rattus norvegicus (strain Wistar) (clone: rpt-1) small intestine
ORGANISM     Rattus norvegicus
              Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.
REFERENCE    1 (bases 1 to 1733)
AUTHORS     Erickson, R.H., Gun, J.R. Jr., Lindstrom, M.M., McKean, D. and Kim, Y.S.
TITLE       Regional expression and dietary regulation of rat small intestinal
              peptide and amino acid transporter mRNAs
JOURNAL      Biochem. Biophys. Res. Commun. 216 (1), 249-257 (1995)
MEDLINE     96067558
PUBMED      7488096
FEATURES
  source     Location/Qualifiers
            1..1733
              /organism="Rattus norvegicus"
              /strain="Wistar"
              /db_xref="taxon:10116"
              /clone="rpt-1"
              /cfeature_type="small intestine"
            <1..>1733
              /codon_start=1
              /product="proton-driven peptide transporter"
              /protein_id="AAC42077.1"
              /db_xref="GI:1082043"
            /translation="LTGRTDGLIKLEKDGILYHARVLDRETRAVHHLRLAALDSGAI
              /VGPVPIILIEVDINDNRPFLQTKRESSVRKGRKFFMYNATLDDLPNQL
              /FOIVQLPKINWVFOIDNKGALISPEBSQVADPKNPYYNLVSVKDWGGNE
              /NSFSOTSDVIDIVRENIMKRAPEVEIREMLTPHPKIKIOVONNEPVAHSLINKEKL
              /POPFSLIDEGNIVYTOPDRBKSHVEFPAKDNKPLAVLEIRYKVIDINDP
              /PTLSQVTVFVQNEENVLSIGIFPAHMDENANNISILMYRLVQTPKVPDELEFL
              /IDYGRKVLGKRSLEKQDSPOYNLTWSDIDFKTCSLOVVIDINQIPIFEESD
              /YSGKLTSEDPAIGSTLILIIQATDDSPFTGSSKILIKYIQGTREGLEAVTDMPTNG
              /YVAKRPDLPEFETEPVSIYFKAEPEPLVNGILEYNASSFASEFLTVXDVEVVFPOQ
              /IFQANSEDTAIGTKVGTATARDPEGLTVSYSLKDKRMLKIDSVTGDIFSTAPLDR
              ETESYRVQVVA"
BASE COUNT   525 a      398 c      393 g      416 t      1 others
ORIGIN
Alignment Scores:
Pred. No.:      9.94e-168      Length:      1733
Score:          2328.00      Matches:      446
Percent Similarity: 86.83%      Conserved: 55
Best Local Similarity: 77.30%      Mismatches: 76
Query Match:    53.88%      Indels: 0
DB:             10      Gaps: 0
US-10-025-380-1081 (1-832) x RATDPRPT (1-1733)
Qy      62  LeuThrGlyGluThrAspAsnIlePheValIleGluArgGluGlyLeuLeuTyrTyrAsn 81
Db      1  CTAACAGAGAAAGCGATGTATATCTTAAGATGAAGAAAGATGACTCTTTATACAGCA 60
Qy      82  ArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAspAla 101
Db      61  AGAGTCTGACAGAGAAACAGAGGCGCTTCATCATCTGGCGCTTGACGCTTGATTC 120
Qy      102  AsnGlyIleIleValGluGlyProValProIleThrIleGluValIleAspIleAsnAsp 121
Db      121  CAGGAGCTATAGTGGATGGCCAGTCCCATCATCATAGAGTCAAGACATCAATGAC 180
Qy      122  AsnArgProThrPheLeuGlnSerIleTyrGlyGlySerValArgGlnAsnSerArgPro 141
Db      181  AACCGACTTCATTTCTCCAGCAAAATATAGAAAGCTCAGTGAAGCGAAGACTCTCGCCA 240
Qy      142  GlyIleProPheLeuTyrValAlaAsnAlaThrAspLeuAspAspProAlaThrProAsnGly 161
Db      241  GGAAGCTTTTCATGATGCTCATGCTACAGACTGATATCCCGCTACTCCCAATGCC 300
Qy      162  GlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPheGln 181

```

```

Db      301 CAGCTTTTATCAAAATGTCATCCAGCTCCCAAAATCAAAATGTTATGATCTTCAG 360
Qy      182 11leasnAnlyThrGlyAlaIleSerLeuThrArgGluSerGlnGluLeuAsnPro 201
Db      361 ATGCAACAACAAGAGGGGCAATTCGCTTACCCCAAGAGTTTCCAGATTTGATCCA 420
Qy      202 AlAlAsnProSerTyrrAsnLeuValIleSerValIlyAspMetGlyGlnSerGlu 221
Db      421 ATAAAGATCCTTACTACAACTGTGTGTCTCTGTGAAGGACATGGGGGGCAGAAATGAG 480
Qy      222 AsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTrrpAla 241
Db      481 AATTCCTTCAGCGATACCATCTGTGAATAATTACTCCAGAGAAACATCTGAAAGCA 540
Qy      242 ProlyProValGluMetValGluAsnSerThrAspProHisProIleValIleThrGln 261
Db      541 CCAGAGCTGTGGAGATTCGAGAAATCTTAATCTGATCTCCACCAATCAAAATCTACAG 600
Qy      262 ValArgTrpAsnAspProGlyAlaGlnTyrrSerLeuValAspLysGluLysLeuProArg 281
Db      601 GTGAGTGAATGAATCAACAGTGGCCCACTATTCCTTATCAACAAGAGAAAGCTGCCAG 660
Qy      282 PheProPheSerIleAspGlnGluGlyAspIleTyrrValThrGlnProLeuAspArgGlu 301
Db      661 TTTCATTCCTCGATCGACCAAGAAATATTAATTCGATCGATCCCTTGGACCGTGAG 720
Qy      302 GluLysAspAlaTyrrValPheTyrrAlaValAlAlAsnAspGluTyrrGlyLysProLeuSer 321
Db      721 GAAAGAGACTCTCATGTGTTTCTTTTCCCACTGCCAAGAGCAAGAAATGSAAAACCCCTTGCA 780
Qy      322 TyrProLeuGlnIleHisValIlyValIlyAspIleAsnAspAsnProProThrCysPro 341
Db      781 TATCACTGGAATTCGTGAAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
Qy      342 SerProValThrValPheGluValGlnGluAsnGluArgLeuGlnIleGlyThr 361
Db      841 TCTCAAGTGAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
Qy      362 LeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrrArgIle 381
Db      901 TTTCGCTGCCCAAGATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 960
Qy      382 ValGluGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrrAlaGly 401
Db      961 GTAGACCAACAACCAAGTTCCTCCCTCAGATGAACTCTTCTCATTTGACGAATATGAGAGA 1020
Qy      402 MetLeuGlnLeuAlAlAsnSerLeuLysGlnAspThrProGlnTyrrAsnLeuThr 421
Db      1021 AAGGTCCAGTTGGGTAAACGCTCTTGAAAGACAGACAGTCCCTCAGTACAACTTAACA 1080
Qy      422 IleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIleAsp 441
Db      1081 GTAGAGGTGTCTGACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140
Qy      442 11leasnAspGlnIleProIlePheGluLysSerAspTyrrGlyAsnLeuThrLeuAlaGlu 461
Db      1141 ATCAACGATCAGATTCCTCCATCTTTGAAGATCAGACTATGGAAGCAAGACTCTTTCGAA 1200
Qy      462 AspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluProPhe 481
Db      1201 GACACAGCCATCCGATCCACATCTTAATCATCTCAAGCTACTGATGATGATGAGCCCTTT 1260
Qy      482 ThrGlySerSerLysIleLeuTyrrHisIleIleLeuGlyAspSerGluArgLeuGly 501
Db      1261 ACCGGAGCTCTTAATTTCTGTACAGATGTACAGGAGACACTGAGGGAAGACTGGA 1320
Qy      502 ValAspThrAspProHisThrAsnThrGlyTyrrValIleIleLysLysProLeuAspPhe 521
Db      1321 GTTGTCACAGATCCCATGACCAATACAGATACGTCACAGATCAGAAAGCCTCTTGATTTT 1380
Qy      522 GluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuValPhe 541
Db      1381 GAAACCGAACCAAGTACTACATGTGTTCAAAGCAGAAAAACCTGAGCGCGTGTAAT 1440

```

Search completed: June 22, 2003, 02:41:16
 Job time : 4969 secs

```

Qy      542 GlyValIleTyrrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspValAsn 561
Db      1441 GGCATAGAGTACAAATGCTAGTCTTTGCTCTCTCTGAGCTGACCTGGAGGATGTGAAT 1500
Qy      562 GluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAlaIle 581
Db      1501 GAAGTCCGTGATTTTCCCGCAAAATATTCCAAGCCATGCTTGGCAGGATACGCTATA 1560
Qy      582 GlyThrLysValGlnLysAsnValThrAlaLysAspProGluGlyLeuAspIleSerTyrrSer 601
Db      1561 GGCATTAAGTGGGCACTGTGACTGCCAGAGATCCGAAGGCTGTGAGTTATTC 1620
Qy      602 LeuArgGlyAspThrArgGlyTrpLeuLysIleAspHisValThrGlyGluIlePheSer 621
Db      1621 CTGAAGACAAATAGAGAGGTTGGCTTAATATCACTCCGTCAACCGGCAATATTTAGC 1680
Qy      622 ValAlaProLeuAspArgGluAlaGlySerProTyrrArgValGlnValVal 638
Db      1681 ACTGCTCCGCTGACAGGGAAACAGAAAGTGTATCGAGTACAGATGCGTG 1731

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2003, 20:20:34 ; Search time 699 Seconds
(without alignments)
1076.720 Million cell updates/sec

Title: US-10-025-380-1076

Perfect score: 3345
Sequence: 1 gaattccgctcgcacacg.....cctttgcgagcgacgaattc 3345

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: N_Geneseq.101002.*
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
25: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3345	100.0	3345	20	AA18166
2	3345	100.0	3345	22	AA129511
3	3345	100.0	3345	24	ABN96791
4	3345	100.0	3345	24	ABL62444
5	3345	100.0	3345	24	ABL64424
6	3314.8	99.1	3654	24	ABN97312
7	3314.8	99.1	3654	24	ABL62484
8	3314.8	99.1	3654	24	ABL67667
9	514.2	15.4	523	24	ABL37430

10	511.4	15.3	627	21	AAA16509	Human colon cancer
11	497.2	14.9	501	24	ABL38503	Human colon tumour
12	469.4	14.0	495	24	ABL36982	Human colon tumour
13	461.6	13.8	492	24	ABL36509	Human colon tumour
14	431.4	12.9	516	24	ABR44858	CDNA encoding colo
15	410.4	12.3	568	24	ABR45154	CDNA encoding colo
16	365.6	10.9	370	22	AAH36589	Human colon tumour
17	360.4	10.8	820	24	AAH33661	Human colon cancer
18	342	10.2	342	24	ABR45744	CDNA encoding colo
19	340.4	10.2	500	24	ABR45083	CDNA encoding colo
20	331.8	9.9	335	23	AA57401	CDNA #77 encoding
21	331.8	9.9	335	23	AA58866	CDNA #1542 encoding
22	331.4	9.9	346	21	AA128480	CDNA encoding huma
23	331.4	9.9	346	22	AA128480	Colon tumour relat
24	319.8	9.6	325	22	AA129276	Colon tumour relat
25	284.8	8.5	591	21	AA16295	Human colon cancer
26	269.6	8.1	274	24	ABL36483	Human colon tumour
27	265.2	7.9	268	24	ABR36452	Human colon cancer
28	265	7.9	265	24	ABR27599	Stomach cancer rel
29	207.8	6.2	167343	24	ABR64403	Thyroid cancer rel
30	207.8	6.2	167343	24	ABR67239	CDNA encoding huma
31	203.8	6.1	3121	24	ABR51416	Human nervous syst
32	199.6	6.0	6618	22	ABR18101	Human nervous syst
33	199.2	6.0	10614	22	AA136948	Human musculocele
34	199	5.9	7582	22	AA164000	Human polynucleoti
35	199	5.9	7582	22	AA331678	Genomic sequence #
36	199	5.9	147419	24	ABR83574	Human CDNA differe
37	197.8	5.9	18061	24	ABR61973	Colon adenocarcino
38	197.8	5.9	18061	24	ABR62480	Colon adenocarcino
39	197.8	5.9	18061	24	ABR62481	Human CDNA sequenc
40	197.4	5.9	3417	22	AAH18467	Human musculocele
41	197.4	5.9	10612	22	AA136946	Human nervous syst
42	196.8	5.9	5161	22	ABR20105	Human prostate exp
43	195.8	5.9	6582	23	ABV50892	Human nervous syst
44	195.8	5.9	6582	22	ABR19716	Human nervous syst
45	195.8	5.9	32169	23	ABR43076	Genomic sequence #

ALIGNMENTS

RESULT 1	
AA18166	
ID	AA18166 standard; DNA; 3345 BP.
AC	AA18166;
DT	04-MAY-1999 (first entry)
XX	
DE	Human HPT-1 protein coding sequence.
XX	
KW	Gastro-intestinal transport receptor; binding protein; hst; HPT1;
KW	D2H; hPPT1; human; GI tract receptor; sucrose-isomaltase complex;
KW	intestinal peptide-associated transporter; hypertension; diabetes;
KW	osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
KW	therapeutic agent delivery; therapy; probe; ss.
OS	Homo sapiens.
XX	
PN	WO9851325-A2.
XX	
PD	19-NOV-1998.
XX	
PF	15-MAY-1998; 98WO-US10088.
XX	
PR	15-MAY-1997; 97US-0046595.
XX	
PA	(CYTO-) CYTOGEN CORP.
XX	(ELAN-) ELAN CORP PLC.
XX	
PI	Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin JV;
XX	Omahony DJ, Patterson CA, Singleton J;
XX	

DR WPI: 1999-009568/01.
 DR P-PSDB; AAW74089.
 XX New proteins that bind specifically to receptors in the
 PT gastro-intestinal tract and related nucleic acid - chimeras and
 PT antibodies, used to deliver therapeutic or diagnostic agents to, or
 PT through, the gastrointestinal tract, e.g. insulin or leuprolide
 PS Disclosure; Fig 2; 294pp; English.
 XX
 CC This sequence encodes the human HPT-1 protein. The invention relates to
 CC purified proteins (I) that bind specifically to at least one of the
 CC gastro-intestinal (GI) tract receptors human intestinal
 CC peptide-associated transporter (HPT1), hPEPT1, D2H and human
 CC sucrose-isomerase complex (HSI). (I) provide active transport of
 CC therapeutic agents through human and animal GI tissue (into the blood)
 CC for in vivo delivery, particularly for treatment or prevention
 CC of hyperension, diabetes, osteoporosis, haemophilia, anaemia, cancer,
 CC migraine, or angina pectoris. Specifically they are used to deliver
 CC insulin or leuprolide, but many other suitable therapeutic agents are
 CC disclosed, including genes or inhibitory nucleic acid, imaging agents and
 CC antigens. (I) may also provide targeting to the GI tract. Other uses of
 CC (I) are: (i) to determine the level of specified receptors in a sample
 CC (in a binding assay); and (ii) to screen for molecules that bind (I).
 CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
 CC and in immunoassays. The antibodies are used to locate, detect and
 CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis
 CC etc., also for peptide purification and immobilisation.
 CC
 XX
 SQ Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;
 Query Match 100.0%; Score 3345; DB 20; Length 3345;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GAATTCGGTCTGACCACTGATGAGAAAGAAAGACTTTTAACCACTTTTGTGACTTA	60
Db	1	GAATTCGGTCTGACCACTGATGAGAAAGAAAGACTTTTAACCACTTTTGTGACTTA	60
Qy	61	CAGAAAGAAATTTGATTAAGAAAAGTATGATCTTCAGGCCATCTTCACTCCCTGTGT	120
Db	61	CAGAAAGAAATTTGATTAAGAAAAGTATGATCTTCAGGCCATCTTCACTCCCTGTGT	120
Qy	121	CTTCTTATGCTTTATTTGGCACTGATATGCGCAAGGGGAAATTATAGGACCCCTG	180
Db	121	CTTCTTATGCTTTATTTGGCACTGATATGCGCAAGGGGAAATTATAGGACCCCTG	180
Qy	181	AAACCCATGACATTTTCTATTATTAAGAGCCAGACGAGTCAAAATTAATTCAGATT	240
Db	181	AAACCCATGACATTTTCTATTATTAAGAGCCAGACGAGTCAAAATTAATTCAGATT	240
Qy	241	AAGCCCAATCTCTCTGCTGTGACTTTGAATACTGGGAGACAGACATATTTGTG	300
Db	241	AAGCCCAATCTCTCTGCTGTGACTTTGAATACTGGGAGACAGACATATTTGTG	300
Qy	301	ATAGAAGGGAGGAGCTTCTGTATTTACAAGAGCTTTGACAGGGAAACAAGATTACT	360
Db	301	ATAGAAGGGAGGAGCTTCTGTATTTACAAGAGCTTTGACAGGGAAACAAGATTACT	360
Qy	361	CACAAATCTCCAGGTTGACAGCCCTGAGCCTAATGAAATTAATGAGGGTCCAGTCCCT	420
Db	361	CACAAATCTCCAGGTTGACAGCCCTGAGCCTAATGAAATTAATGAGGGTCCAGTCCCT	420
Qy	421	ATCAACATAGAAGTGAAGACATCAACGACATGACCCAGTTTCTCACTCAAAAGTAC	480
Db	421	ATCAACATAGAAGTGAAGACATCAACGACATGACCCAGTTTCTCACTCAAAAGTAC	480
Qy	481	GAAAGCTCAATGAAGCAGAACTCTGCCCCAGAAAGCCCTTCTGTATGTAAGGCCACA	540
Db	481	GAAAGCTCAATGAAGCAGAACTCTGCCCCAGAAAGCCCTTCTGTATGTAAGGCCACA	540
Qy	541	GACCTGATGATCGGGCACTCCCAATGGCCAGCTTTATTAACGATTTGATCCAGCTT	600
Db	541	GACCTGATGATCGGGCACTCCCAATGGCCAGCTTTATTAACGATTTGATCCAGCTT	600
Qy	601	CCCATGATCAACATATGATGATCTTTGATGATTAACAAGAAACGGAGCATCTCTCT	660
Db	601	CCCATGATCAACATATGATGATCTTTGATGATTAACAAGAAACGGAGCATCTCTCT	660
Qy	661	ACCCGAGAGGATCTCAGGAAATGGAATTCCTGTAGAAATCTCTTAATCTGTGATC	720
Db	661	ACCCGAGAGGATCTCAGGAAATGGAATTCCTGTGTAAATCTCTTAATCTGTGATC	720
Qy	721	TCAGTGAAGACATGAGGAGCCAGAGTGAAGATTCCTTCAGTATACCATCTGTGAT	780
Db	721	TCAGTGAAGACATGAGGAGCCAGAGTGAAGATTCCTTCAGTATACCATCTGTGAT	780
Qy	781	ATCATATGACACAGAAATTTTGAAGACACCAAAACCTGTGAGATGAGGAAACTCA	840
Db	781	ATCATATGACACAGAAATTTTGAAGACACCAAAACCTGTGAGATGAGGAAACTCA	840
Qy	841	ACTGATCTCACCCTCAAAATCACTCAGTGCAGTGTGATGATCCCGTGCACAAATAT	900
Db	841	ACTGATCTCACCCTCAAAATCACTCAGTGCAGTGTGATGATCCCGTGCACAAATAT	900
Qy	901	TCCTTAGTGAACAAGAAAGCTGCCAAGATTCCTCAATTTCAATTAACAGAAAGAT	960
Db	901	TCCTTAGTGAACAAGAAAGCTGCCAAGATTCCTCAATTTCAATTAACAGAAAGAT	960
Qy	961	ATTATGCTGACTCAGCCCTTTGACCGAGAAAGAAAGATGATGTTTATGACGTT	1020
Db	961	ATTATGCTGACTCAGCCCTTTGACCGAGAAAGAAAGATGATGTTTATGACGTT	1020
Qy	1021	GCAAAGATGATGATGAGAAAGCACTTTCATATCCCTGGAATTCATGTAAGTTAA	1080
Db	1021	GCAAAGATGATGATGAGAAAGCACTTTCATATCCCTGGAATTCATGTAAGTTAA	1080
Qy	1081	GATATTAATGATATATCCACTACATGTCGTGACACAGTAAACGTAATTTGAGTCCAGAG	1140
Db	1081	GATATTAATGATATATCCACTACATGTCGTGACACAGTAAACGTAATTTGAGTCCAGAG	1140
Qy	1141	AATGAAGATCTGGTAAACAGTATTCGGGACCCCTTAACAGACATGAGGATGAAAGAAAT	1200
Db	1141	AATGAAGATCTGGTAAACAGTATTCGGGACCCCTTAACAGACATGAGGATGAAAGAAAT	1200
Qy	1201	ACTGCAACAGTTTCTTAACTACAGATTTGAGACAAACCTCCAACTTCCCATGAT	1260
Db	1201	ACTGCAACAGTTTCTTAACTACAGATTTGAGACAAACCTCCAACTTCCCATGAT	1260
Qy	1261	GGAATCTTCTTAATCCAAACCTATGCTGAAATGTAACGTTAAGCTCTTGAAG	1320
Db	1261	GGAATCTTCTTAATCCAAACCTATGCTGAAATGTAACGTTAAGCTCTTGAAG	1320
Qy	1321	AAGCAAGATCTCTCTCAAGTCAACTTAACGATAGAGTGTGACAAAGATTCAAGACC	1380
Db	1321	AAGCAAGATCTCTCTCAAGTCAACTTAACGATAGAGTGTGACAAAGATTCAAGACC	1380
Qy	1381	CTTTGTTTGTGCAAAATCAACGTTAATGATATCAATGATCCCATCTTTGAAAAA	1440
Db	1381	CTTTGTTTGTGCAAAATCAACGTTAATGATATCAATGATCCCATCTTTGAAAAA	1440
Qy	1441	TCAGATTAATGAAACCTGACTCTTGTGAAAGACAAACATTTGGGTCCACCATCTTAAC	1500
Db	1441	TCAGATTAATGAAACCTGACTCTTGTGAAAGACAAACATTTGGGTCCACCATCTTAAC	1500
Qy	1501	ATCCAGGCTCAGATGCTGATGAGCACTTTAATGAGGAGTTCTTAATCTGTATCATATC	1560
Db	1501	ATCCAGGCTCAGATGCTGATGAGCACTTTAATGAGGAGTTCTTAATCTGTATCATATC	1560
Qy	1561	ATAAAGGAGACAGTGAAGGAGCCCTGGGGGTTGACACAGATCCCATCAACCCGGA	1620
Db	1561	ATAAAGGAGACAGTGAAGGAGCCCTGGGGGTTGACACAGATCCCATCAACCCGGA	1620
Qy	1621	TATGTCAATTAATTAAGCTCTTGAATTTTGAACAGACAGCTGTTCATCATTTGTTC	1680
Db	1621	TATGTCAATTAATTAAGCTCTTGAATTTTGAACAGACAGCTGTTCATCATTTGTTC	1680

QY 1681 AAAGCAGAAAATCTGAGCCCTAGTGTGTTGGTGAAGTCAATGCAAGTTCTTTGGC 1740
 DB 1681 AAAGCAGAAAATCTGAGCCCTAGTGTGTTGGTGAAGTCAATGCAAGTTCTTTGGC 1740
 QY 1741 AAGTTCACGCTTATTTGACAGATGTGAAGACCTCAATTTTCCCAACAGTATTC 1800
 DB 1741 AAGTTCACGCTTATTTGACAGATGTGAAGACCTCAATTTTCCCAACAGTATTC 1800
 QY 1801 CAAGCGAAGTCAGTGAAGATGTAGCTATGAGGCACTAAAGGGCAGATGACGCCAAG 1860
 DB 1801 CAAGCGAAGTCAGTGAAGATGTAGCTATGAGGCACTAAAGGGCAGATGACGCCAAG 1860
 QY 1861 GATCCAGAAGGTCGTGACATTAAGCTATTTCACTGAGGGAGACACAAAGGTTGGCTTAA 1920
 DB 1861 GATCCAGAAGGTCGTGACATTAAGCTATTTCACTGAGGGAGACACAAAGGTTGGCTTAA 1920
 QY 1921 ATTGACCAAGTACTGAGTGAAGATCTTTAGTGTGCTCCATTGGAACAGAAAGCCGGAAGT 1980
 DB 1921 ATTGACCAAGTACTGAGTGAAGATCTTTAGTGTGCTCCATTGGAACAGAAAGCCGGAAGT 1980
 QY 1981 CCAATATGGGATCAAGTGTGAGCCACAGAAAGTGAAGGGGGTCTTCCCTTAAGCTCTGTGCA 2040
 DB 1981 CCAATATGGGATCAAGTGTGAGCCACAGAAAGTGAAGGGGGTCTTCCCTTAAGCTCTGTGCA 2040
 QY 2041 GAGTTCACCTGATCCTTAATGATGATGATGACAAACCTCCAGGCTAGCCAGAGACTAC 2100
 DB 2041 GAGTTCACCTGATCCTTAATGATGATGATGACAAACCTCCAGGCTAGCCAGAGACTAC 2100
 QY 2101 ACGGGCTGTGTTCTTCCCATCCCTCAGTGCACCTGGAAGTCTCATTTTCAGAGGCTACT 2160
 DB 2101 ACGGGCTGTGTTCTTCCCATCCCTCAGTGCACCTGGAAGTCTCATTTTCAGAGGCTACT 2160
 QY 2161 GATGATGATGACATTAATTCGGGGTCCCATTTTCAATTTTCCCTCCGAGTGGAAAGC 2220
 DB 2161 GATGATGATGACATTAATTCGGGGTCCCATTTTCAATTTTCCCTCCGAGTGGAAAGC 2220
 QY 2221 TTACAAAACGACTGGGAAGTTTCCAAAATCATAGTGTCTCATGCGCAGTGTCTACAGG 2280
 DB 2221 TTACAAAACGACTGGGAAGTTTCCAAAATCATAGTGTCTCATGCGCAGTGTCTACAGG 2280
 QY 2281 CACACAGACTTTGAGAGAGAGGGCGATGTGCTTGAATCCGACATGATGATGGGGGTCG 2340
 DB 2281 CACACAGACTTTGAGAGAGAGGGCGATGTGCTTGAATCCGACATGATGATGGGGGTCG 2340
 QY 2341 CCAACCTTGAAGGATTTGTTTCTTACAGTTACATTTGCAAGTGTGTGGAAGGAAGT 2400
 DB 2341 CCAACCTTGAAGGATTTGTTTCTTACAGTTACATTTGCAAGTGTGTGGAAGGAAGT 2400
 QY 2401 TGTTCGAGGCGACAGGTCACAGACTGGGATCCCACTGAGGCGATGGAGTTGTATTA 2460
 DB 2401 TGTTCGAGGCGACAGGTCACAGACTGGGATCCCACTGAGGCGATGGAGTTGTATTA 2460
 QY 2461 CTGCTGACCAACCTTCTGTGTGTATTAATTTTACAGTGTGTATTCGCCATTAAG 2520
 DB 2461 CTGCTGACCAACCTTCTGTGTGTATTAATTTTACAGTGTGTATTCGCCATTAAG 2520
 QY 2521 AAGGATTAAGGCAAAATATATTTGAAAGTGTCAAGCATCTGAAGTCAAACTCTGAGA 2580
 DB 2521 AAGGATTAAGGCAAAATATATTTGAAAGTGTCAAGCATCTGAAGTCAAACTCTGAGA 2580
 QY 2581 AGCTGAATTTGAAGAAGATGTTGTAATTTATATAGCAAGTGTGATTTGCAACAACA 2640
 DB 2581 AGCTGAATTTGAAGAAGATGTTGTAATTTATATAGCAAGTGTGATTTGCAACAACA 2640
 QY 2641 TCTCATCTTATTAATTTTCTTATCATCAAGTGTATTAATTTTAAACAGATATTCCTCT 2700
 DB 2641 TCTCATCTTATTAATTTTCTTATCATCAAGTGTATTAATTTTAAACAGATATTCCTCT 2700
 QY 2701 TGTCTTTAATATTTGTAATATTTCTTTTGAAGTGAAGTGTGCTGTGCGCCAG 2760
 DB 2701 TGTCTTTAATATTTGTAATATTTCTTTTGAAGTGAAGTGTGCTGTGCGCCAG 2760

QY 2761 GCTGAGATCAGATGTGTATGCCAGCTCACTGCAACCTTCGCTCGGATTCACATGA 2820
 DB 2761 GCTGAGATCAGATGTGTATGCCAGCTCACTGCAACCTTCGCTCGGATTCACATGA 2820
 QY 2821 TTCTCTGCTCAGCTTCTTAAGTGTGGTTTACAGGACCCACCATGCTCCAGCT 2880
 DB 2821 TTCTCTGCTCAGCTTCTTAAGTGTGGTTTACAGGACCCACCATGCTCCAGCT 2880
 QY 2881 AATTTTGTATTTTAAATAGACGAGGCTTTCGCCATTTGGCAGAGTGTGCTGAATC 2940
 DB 2881 AATTTTGTATTTTAAATAGACGAGGCTTTCGCCATTTGGCAGAGTGTGCTGAATC 2940
 QY 2941 CTGACGTCAGATGATGCTGCTGCTGTGTCTCCCAATACAGGATGACCATGCCA 3000
 DB 2941 CTGACGTCAGATGATGCTGCTGCTGTGTCTCCCAATACAGGATGACCATGCCA 3000
 QY 3001 CTTACTAGATATTTCAATGTGCTATAGACATTAGACAGATTTTTCATTTTCAATGACAT 3060
 DB 3001 CTTACTAGATATTTCAATGTGCTATAGACATTAGACAGATTTTTCATTTTCAATGACAT 3060
 QY 3061 TTTTCCCTCTGCAAAATGAGCTTACTGCTGCTTTTCCCTTTGGGCGCAACAGACT 3120
 DB 3061 TTTTCCCTCTGCAAAATGAGCTTACTGCTGCTTTTCCCTTTGGGCGCAACAGACT 3120
 QY 3121 CATTAAATATTTCTGATCAATTTTCTTTATCAAGAGATATACAGTGTGCTCATAGA 3180
 DB 3121 CATTAAATATTTCTGATCAATTTTCTTTATCAAGAGATATACAGTGTGCTCATAGA 3180
 QY 3181 ACTGCTGATTCATTTATGTTTTTGTGATTCATCTGTGCTCCCTTTCATCTTGAC 3240
 DB 3181 ACTGCTGATTCATTTATGTTTTTGTGATTCATCTGTGCTCCCTTTCATCTTGAC 3240
 QY 3241 TCCCTTGTATTTTCACTGATTTTCAAAATTTGTCTGAGAGAGAGAGAGAGAGACT 3300
 DB 3241 TCCCTTGTATTTTCACTGATTTTCAAAATTTGTCTGAGAGAGAGAGAGAGAGACT 3300
 QY 3301 GAAAAATTAATTAATTAAGAACAGCCTTTTGGGCGCGCAATTC 3345
 DB 3301 GAAAAATTAATTAATTAAGAACAGCCTTTTGGGCGCGCAATTC 3345

RESULT 2
 AAI29511
 ID AAI29511 standard; cDNA; 3345 BP.
 XX
 AC AAI29511;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Human intestinal peptide-associated transporter HPT-1 cDNA sequence.
 XX
 KW Human; immunotherapy; diagnosis; colon cancer; colon tumour;
 KW immunogenic; gene therapy; vaccine; colonic cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200149716-A2.
 XX
 PD 12-JUL-2001.
 XX
 PF 29-DEC-2000; 2000MO-US35596.
 XX
 PR 30-DEC-1999; 99US-0476296.
 PR 10-JAN-2000; 2000US-0480321.
 PR 15-FEB-2000; 2000US-0504629.
 PR 06-MAR-2000; 2000US-0519444.
 PR 19-MAY-2000; 2000US-0575251.
 PR 29-JUN-2000; 2000US-0609448.
 PR 28-AUG-2000; 2000US-0649811.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 XU Xu J, Lodes MJ, Secret H, Benson DR, Meagher MJ, Stolk JA;

PI King GE, Wang T, Jiang Y;

XX WPI; 2001-441847/47.

XX Colon tumor associated proteins and nucleic acids useful for the
PT prevention, diagnosis and treatment of colonic cancer -
XX

PS Claim 2; Page 466-467; 472pp; English.

XX The present invention describes colon tumour associated proteins (I) and
CC the polynucleotides (II) that encode them. (II) have cytostatic activity.
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
CC (II) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate colon tumour associated protein (TCAP)
CC expression, such as colonic cancer. For example, (I) and (II) may be
CC used to treat disorders associated with decreased expression by
CC rectifying mutations or deletions in a patient's genome that affect the
CC activity of TCAPs by expressing inactive proteins or to supplement the
CC patients own production of them. Additionally, (II) may be used to
CC produce the TCAP proteins, by inserting the nucleic acids into a host
CC cell culturing the cell to express the protein. (II) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC polymerase chain reaction (PCR) and hybridisation assays to detect and
CC quantitate the presence of similar nucleic acids in samples, and (I) may
CC therefore be used as antigens in the production of antibodies against TCAPs
CC and in assays to identify modulators of TCAP expression and activity.
CC Anti-(I) antibodies and antagonists may also be used to down regulate
CC TCAP expression and activity. The anti-(I) antibodies may also be used
CC as diagnostic agents for detecting the presence of TCAPs in samples
CC (e.g. by enzyme linked immunosorbent assay (ELISA)). AA128460 to AA129512
CC and AA124494 to AA124523 represent nucleotide and amino acid sequences
CC given in the exemplification of the present invention.

XX Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;

SQ Query Match 100.0%; Score 3345; DB 22; Length 3345;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGGCTCGACACTGATGAGAAAGAAAGAACTTTTAAACCACTTTTGACTTA 60
DB 1 GAATTCGGCTCGACACTGATGAGAAAGAAAGAAAGAACTTTTAAACCACTTTTGACTTA 60
QY 61 CAGAAAGAAATTTGATTAAGAAACCTATGATCTTCAAGGCCCATCTTCCCTGTGT 120
DB 61 CAGAAAGAAATTTGATTAAGAAACCTATGATCTTCAAGGCCCATCTTCCCTGTGT 120
QY 121 CTTCTTATGCTTTATTTGGCAACTGATATGAGCCAAAGAGGAAATTAGTGACCCCTG 180
DB 121 CTTCTTATGCTTTATTTGGCAACTGATATGAGCCAAAGAGGAAATTAGTGACCCCTG 180
QY 181 AAACCCATGACATTTTCTATTATTAAGAGCCAAAGACCGAGTCAAAATTATATCCAGTTT 240
DB 181 AAACCCATGACATTTTCTATTATTAAGAGCCAAAGACCGAGTCAAAATTATATCCAGTTT 240
QY 241 AAGGCCAATCTCTCTGCTGTGACTTTTGAATTAAGTGGGAGAGACAAATATTTGTG 300
DB 241 AAGGCCAATCTCTCTGCTGTGACTTTTGAATTAAGTGGGAGAGACAAATATTTGTG 300
QY 301 ATAGAAGGGAGGGAGCTCTGTATTTACAAGAGGCTTGGAAGGAAACAAAGATCTACT 360
DB 301 ATAGAAGGGAGGGAGCTCTGTATTTACAAGAGGCTTGGAAGGAAACAAAGATCTACT 360
QY 361 CACAATCTCCAGGTTCAGCCCTGAGCGCTAATGGAATTAATAGTGAAGGTCCAGTCCCT 420
DB 361 CACAATCTCCAGGTTCAGCCCTGAGCGCTAATGGAATTAATAGTGAAGGTCCAGTCCCT 420
QY 421 ATCAACATAGAAGTGAAGAGACATCAAGACATGAGACCAAGCTTTTCTCAATCAAAGTAC 480
DB 421 ATCAACATAGAAGTGAAGAGACATCAAGACATGAGACCAAGCTTTTCTCAATCAAAGTAC 480
QY 481 GAAGGCTCAGTAAGGACGAAGCTCTGCCAAGAAAGCCCTTCTGTATGTCAATGCCACA 540

DB 481 GAAGGCTCAGTAAGGACGAAGCTCTGCCAAGAAAGCCCTTCTGTATGTCAATGCCACA 540
QY 541 GACCTGATGATCCGGCCCACTCCCAATGGCCAGCTTTATTAACAGATTGCAATCCAGCTT 600
DB 541 GACCTGATGATCCGGCCCACTCCCAATGGCCAGCTTTATTAACAGATTGCAATCCAGCTT 600
QY 601 CCATGATCAACAATGCTCATGTACTTTCAGATCAACAACAAAGGAGGCATCTCTCT 660
DB 601 CCATGATCAACAATGCTCATGTACTTTCAGATCAACAACAAAGGAGGCATCTCTCT 660
QY 661 ACCGAGAGGATCTCAGGAATTTGAATTCCTGTAAGATCTTCTTAATCTTGATGATC 720
DB 661 ACCGAGAGGATCTCAGGAATTTGAATTCCTGTAAGATCTTCTTAATCTTGATGATC 720
QY 721 TCAATGAGACATGAGAGGAGCCAGATGAGAAATTCCTTCAAGTATACCAATCTGAGAT 780
DB 721 TCAATGAGACATGAGAGGAGCCAGATGAGAAATTCCTTCAAGTATACCAATCTGAGAT 780
QY 781 ATCATATGTGACAGAAATATTTTGAAGACCAAAACCTGTGAGATGTGAAAATCTCA 840
DB 781 ATCATATGTGACAGAAATATTTTGAAGACCAAAACCTGTGAGATGTGAAAATCTCA 840
QY 841 ACTGATCTTCACCCCATCAAAATCACTCAGGTGCGGTGGAATGATCCGGTGCAATAT 900
DB 841 ACTGATCTTCACCCCATCAAAATCACTCAGGTGCGGTGGAATGATCCGGTGCAATAT 900
QY 901 TCCCTTATGTCACAAAGAAAGAGTCCCAATTTCCATTTTCAATTAACCAAGAAAGAT 960
DB 901 TCCCTTATGTCACAAAGAAAGAGTCCCAATTTCCATTTTCAATTAACCAAGAAAGAT 960
QY 961 ATTTAGTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
DB 961 ATTTAGTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
QY 1021 GCAAGAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
DB 1021 GCAAGAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
QY 1081 GATATTTATATATATCACTACATGTCCTGACCAAGTAACTGATTTGAGGTCAGGAG 1140
DB 1081 GATATTTATATATATCACTACATGTCCTGACCAAGTAACTGATTTGAGGTCAGGAG 1140
QY 1141 AATGAAAGCACTGGGTAACAGTATCGGAGCCCTTATCTGACATGACAGGATGAAAGAAAT 1200
DB 1141 AATGAAAGCACTGGGTAACAGTATCGGAGCCCTTATCTGACATGACAGGATGAAAGAAAT 1200
QY 1201 ACTGCCAAGATTTTCTTAACTACAGAGTGTGAGCAAACTCCCAACTTCCCATGAT 1260
DB 1201 ACTGCCAAGATTTTCTTAACTACAGAGTGTGAGCAAACTCCCAACTTCCCATGAT 1260
QY 1261 GGAATCTTCTTAACTTCAAACTTACAGATGTGAGATGTACAGTAACTGATCTTGAAG 1320
DB 1261 GGAATCTTCTTAACTTCAAACTTACAGATGTGAGATGTACAGTAACTGATCTTGAAG 1320
QY 1321 AAGCAAGATCTCTCTGATCACTTAAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1380
DB 1321 AAGCAAGATCTCTCTGATCACTTAAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1380
QY 1381 CTTTGTGTTGTCAGAACTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1440
DB 1381 CTTTGTGTTGTCAGAACTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1440
QY 1441 TCAATATATGAAACCTGATCTTCTGTAAGACCAAACTTGGGTCCACATCTTAAC 1500
DB 1441 TCAATATATGAAACCTGATCTTCTGTAAGACCAAACTTGGGTCCACATCTTAAC 1500
QY 1501 ATCCAGGCACTGATGCTGATGAGCACTTAATGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
DB 1501 ATCCAGGCACTGATGCTGATGAGCACTTAATGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
QY 1561 ATAAAGGAGACAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620

Db 1561 ATAAAGGAGACAGTGAAGGAGCGCTGGGGGTGACACAGATCCCATACACACCGGA 1620
 Qy 1621 TATGTCATTAATTAATAAAAGCTCTGATTTTGAACAGACAGCTGTTTCCAACTATGTGTC 1680
 Db 1621 TATGTCATTAATTAATAAAAGCTCTGATTTTGAACAGACAGCTGTTTCCAACTATGTGTC 1680
 Qy 1681 AAAGCAGAAAATCTGAGCCTCTAGTGTGTTGGTGAAGTCAATGCAAGTCTTTTGGC 1740
 Db 1681 AAAGCAGAAAATCTGAGCCTCTAGTGTGTTGGTGAAGTCAATGCAAGTCTTTTGGC 1740
 Qy 1741 AAGTTCACGCTTATTTGAGACAGATGTGAATGACACCTCAATTTTCCCAACAGTATTC 1800
 Db 1741 AAGTTCACGCTTATTTGAGACAGATGTGAATGACACCTCAATTTTCCCAACAGTATTC 1800
 Qy 1801 CAAGCAGAAAGTCAGTAGAGATGTAGCTATAGGACTAAAGTGGGCAATGTGACTGCCAAG 1860
 Db 1801 CAAGCAGAAAGTCAGTAGAGATGTAGCTATAGGACTAAAGTGGGCAATGTGACTGCCAAG 1860
 Qy 1861 GATCCAGAAAGTCTGACATTAAGCTATTCATGAGGGGAGACACAGAGAGTGGCTTAA 1920
 Db 1861 GATCCAGAAAGTCTGACATTAAGCTATTCATGAGGGGAGACACAGAGAGTGGCTTAA 1920
 Qy 1921 ATTGACACGCTGATGTGTGAGATCTTTAGTGTGCTCCATTTGACAGAGAACCGGAAGT 1980
 Db 1921 ATTGACACGCTGATGTGTGAGATCTTTAGTGTGCTCCATTTGACAGAGAACCGGAAGT 1980
 Qy 1981 CCATATCGGGTACAAAGTGTGGCCACAGAAAGTAGGGGGGTCTCTTAAAGCTGTGTGCA 2040
 Db 1981 CCATATGGGGTACAAAGTGTGGCCACAGAAAGTAGGGGGGTCTCTTAAAGCTGTGTGCA 2040
 Qy 2041 GAGTTCACCTGATCTTATGATGTGAATGACAAACCTCCAGGCTAGCCAGAGACTAC 2100
 Db 2041 GAGTTCACCTGATCTTATGATGTGAATGACAAACCTCCAGGCTAGCCAGAGACTAC 2100
 Qy 2101 ACGGCTGTTGTTCTTTCGCCATCCCTCAGTGCACCTGGAAGTCTCATTTTCGAGGCTACT 2160
 Db 2101 ACGGCTGTTGTTCTTTCGCCATCCCTCAGTGCACCTGGAAGTCTCATTTTCGAGGCTACT 2160
 Qy 2161 GATGATGATGACACTTATTTCCGGGGTCCCATTTTACATTTTCCCTGGGAGTGGAAAGC 2220
 Db 2161 GATGATGATGACACTTATTTCCGGGGTCCCATTTTACATTTTCCCTGGGAGTGGAAAGC 2220
 Qy 2221 TTACAAAACGACTGGGAAGTTTCCAAAATCAATGTGACTCATGCCGACGTCATCCAGG 2280
 Db 2221 TTACAAAACGACTGGGAAGTTTCCAAAATCAATGTGACTCATGCCGACGTCATCCAGG 2280
 Qy 2281 CACACAGACTTTGAGAGAGAGGGCGTATGTGCTTGAATCCGATCAATGATGGGGGTGCG 2340
 Db 2281 CACACAGACTTTGAGAGAGAGGGCGTATGTGCTTGAATCCGATCAATGATGGGGGTGCG 2340
 Qy 2341 CCAACCTTGGAGGCACTTGTCTTTACAGTTACATTTGCGAGTTGTGTGGAAGAGT 2400
 Db 2341 CCAACCTTGGAGGCACTTGTCTTTACAGTTACATTTGCGAGTTGTGTGGAAGAGT 2400
 Qy 2401 TGTTCGCGGCGAGAGGTCAACAGACTGGATACCACTGTGGCATGGAGTGTGATATA 2460
 Db 2401 TGTTCGCGGCGAGAGGTCAACAGACTGGATACCACTGTGGCATGGAGTGTGATATA 2460
 Qy 2461 CTGCTGACCACTCTTCTGTGTGATGTGTATATTAGCAAGTGTGTTATCCGATTAAG 2520
 Db 2461 CTGCTGACCACTCTTCTGTGTGATGTGTATATTAGCAAGTGTGTTATCCGATTAAG 2520
 Qy 2521 AAGGATTAAGGCAAGATTAATGTTGAAGGTCAAGCACTGAAGTCAAACTCTGAGA 2580
 Db 2521 AAGGATTAAGGCAAGATTAATGTTGAAGGTCTCAAGCACTGAAGTCAAACTCTGAGA 2580
 Qy 2581 AGCTGAATTTGAAAGGAATGTTGAATTTATATAGCAAGTGTGATTTGACAAACCA 2640
 Db 2581 AGCTGAATTTGAAAGGAATGTTGAATTTATATAGCAAGTGTGATTTGACAAACCA 2640
 Qy 2641 TCTCATCTTATTAATTTTCACTAAGTGCATTAATTTTAAACAGATATTCCTCT 2700
 Db 2641 TCTCATCTTATTAATTTTCACTAAGTGCATTAATTTTAAACAGATATTCCTCT 2700

Qy 2701 TGTCTTATTAATTTGCTAAATATTTCTTTTGAAGTGAAGTCTGTCGCGCCAG 2760
 Db 2701 TGTCTTATTAATTTGCTAAATATTTCTTTTGAAGTGAAGTCTGTCGCGCCAG 2760
 Qy 2761 GCTGAGTACAGTGTGTGATGCCAGCTCACTCAACCTCCGCTCTGAGTTTCAATGA 2820
 Db 2761 GCTGAGTACAGTGTGTGATGCCAGCTCACTCAACCTCCGCTCTGAGTTTCAATGA 2820
 Qy 2821 TTCTCTGCTCAGCTTCCTAAGTACTGGGTTTACAGGACCCACCAATGCCAGCT 2880
 Db 2821 TTCTCTGCTCAGCTTCCTAAGTACTGGGTTTACAGGACCCACCAATGCCAGCT 2880
 Qy 2881 AATTTTGTATTTTATATAGACAGGGGTTTCCCAATTTGGCCAGGCTGTGTTGAATC 2940
 Db 2881 AATTTTGTATTTTATATAGACAGGGGTTTCCCAATTTGGCCAGGCTGTGTTGAATC 2940
 Qy 2941 CTGACGTCAGTGTATCTGCTGCTTGTGCTCCCAATACAGGCATGAACCATGCCA 3000
 Db 2941 CTGACGTCAGTGTATCTGCTGCTTGTGCTCCCAATACAGGCATGAACCATGCCA 3000
 Qy 3001 CCTACTAGTATTTTCAATGTGCTATAGCAATTAAGAGATTTTTCATTTTCCATGACAT 3060
 Db 3001 CCTACTAGTATTTTCAATGTGCTATAGCAATTAAGAGATTTTTCATTTTCCATGACAT 3060
 Qy 3061 TTTTCTCTCTGCAAAATGCTTAGCTACTGTGTTTCCCTTTGGGGCAGACAGACT 3120
 Db 3061 TTTTCTCTCTGCAAAATGCTTAGCTACTGTGTTTCCCTTTGGGGCAGACAGACT 3120
 Qy 3121 CATTAAATATCTGTACATTTTCTTTATCAAGAGATATATCAGTGTGTCATAGA 3180
 Db 3121 CATTAAATATCTGTACATTTTCTTTATCAAGAGATATATCAGTGTGTCATAGA 3180
 Qy 3181 ACTGCTGATTCATTTATGTTTTCGATTCATCTGTGTCCCTTATCCTTTGAC 3240
 Db 3181 ACTGCTGATTCATTTATGTTTTCGATTCATCTGTGTCCCTTATCCTTTGAC 3240
 Qy 3241 TCCTTGTGATTTTCACTGATTTCAAAACATTTGTGAGAGAAAAGTGAAGTCTCAG 3300
 Db 3241 TCCTTGTGATTTTCACTGATTTCAAAACATTTGTGAGAGAAAAGTGAAGTCTCAG 3300
 Qy 3301 GAAAAATTAATTAATTAAGAACAGCCTTTTGGCGCGGAATTC 3345
 Db 3301 GAAAAATTAATTAATTAAGAACAGCCTTTTGGCGCGGAATTC 3345

RESULT 3
 AEN96791
 ID AEN96791 standard; DNA; 3345 BP.
 XX
 AC AEN96791;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Gene #13289 used to diagnose liver cancer.
 XX
 KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 KW metastatic liver tumour; cytostatic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 OS Homo sapiens.
 XX
 PN W0200229103-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 02-OCT-2001; 2001WO-US30589.
 XX
 PR 02-OCT-2000; 2000US-237054P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX WPI: 2002-426119/45.
 DR Diagnosing and detecting the progression of liver cancer.
 XX hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample -
 XX Claim 1; SEQ ID NO 3289; 298bp; English.
 XX
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumor in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytosolic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;
 Query Match 100.0%; Score 3345; DB 24; Length 3345;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAATTCGCTCGCACTGATGAGAAAGACCTTTAACACCATTTTGTGACTTA 60
 DB 1 GAATTCGCTCGCACTGATGAGAAAGACCTTTAACACCATTTTGTGACTTA 60
 QY 61 CAGAAAGAAATTTGAATTAAGAAACCTATGATCTTCAGGCCCATCTTCACTCCCTGT 120
 DB 61 CAGAAAGAAATTTGAATTAAGAAACCTATGATCTTCAGGCCCATCTTCACTCCCTGT 120
 QY 121 CTTCTTATGCTTTATTTGGCACTGATATGCGCAAGAGGGAAATTTAGTGAACCCCTG 180
 DB 121 CTTCTTATGCTTTATTTGGCACTGATATGCGCAAGAGGGAAATTTAGTGAACCCCTG 180
 QY 122 CTTCTTATGCTTTATTTGGCACTGATATGCGCAAGAGGGAAATTTAGTGAACCCCTG 180
 DB 122 CTTCTTATGCTTTATTTGGCACTGATATGCGCAAGAGGGAAATTTAGTGAACCCCTG 180
 QY 181 AAACCCATGACATTTTCTATTTATGAAGCCCAAGACCGAGTCAATTTATTTCCAGTT 240
 DB 181 AAACCCATGACATTTTCTATTTATGAAGCCCAAGACCGAGTCAATTTATTTCCAGTT 240
 QY 181 AAACCCATGACATTTTCTATTTATGAAGCCCAAGACCGAGTCAATTTATTTCCAGTT 240
 DB 181 AAACCCATGACATTTTCTATTTATGAAGCCCAAGACCGAGTCAATTTATTTCCAGTT 240
 QY 241 AAGCCCAATCTCTGCTGCTGACTTTTGAATTAATGAGGAGAGACAAATATTTGTG 300
 DB 241 AAGCCCAATCTCTGCTGCTGACTTTTGAATTAATGAGGAGAGACAAATATTTGTG 300
 QY 241 AAGCCCAATCTCTGCTGCTGACTTTTGAATTAATGAGGAGAGACAAATATTTGTG 300
 DB 241 AAGCCCAATCTCTGCTGCTGACTTTTGAATTAATGAGGAGAGACAAATATTTGTG 300
 QY 301 ATAGAACGGAGGAACTTCTGATTTTCAACAGAGCTTTGACAGGGAAACAAATCTTACT 360
 DB 301 ATAGAACGGAGGAACTTCTGATTTTCAACAGAGCTTTGACAGGGAAACAAATCTTACT 360
 QY 301 ATAGAACGGAGGAACTTCTGATTTTCAACAGAGCTTTGACAGGGAAACAAATCTTACT 360
 DB 301 ATAGAACGGAGGAACTTCTGATTTTCAACAGAGCTTTGACAGGGAAACAAATCTTACT 360
 QY 361 CACATATCTCCAGTTTGCAGCCCTGAGCGCTAATGGAATTAATGAGGAGTCCAGTCCCT 420
 DB 361 CACATATCTCCAGTTTGCAGCCCTGAGCGCTAATGGAATTAATGAGGAGTCCAGTCCCT 420
 QY 361 CACATATCTCCAGTTTGCAGCCCTGAGCGCTAATGGAATTAATGAGGAGTCCAGTCCCT 420
 DB 361 CACATATCTCCAGTTTGCAGCCCTGAGCGCTAATGGAATTAATGAGGAGTCCAGTCCCT 420
 QY 421 ATGACCATAGAAAGTGAAGACATCAACAGACATGCAACCGTTTCTCAATGCAAGTAC 480
 DB 421 ATGACCATAGAAAGTGAAGACATCAACAGACATGCAACCGTTTCTCAATGCAAGTAC 480
 QY 421 ATGACCATAGAAAGTGAAGACATCAACAGACATGCAACCGTTTCTCAATGCAAGTAC 480
 DB 421 ATGACCATAGAAAGTGAAGACATCAACAGACATGCAACCGTTTCTCAATGCAAGTAC 480
 QY 481 GAAGGCTCAGTGAAGCAGAACTCTGCGCCAGAAAGCCCTTCTGTATGTCAATGCCA 540
 DB 481 GAAGGCTCAGTGAAGCAGAACTCTGCGCCAGAAAGCCCTTCTGTATGTCAATGCCA 540
 QY 481 GAAGGCTCAGTGAAGCAGAACTCTGCGCCAGAAAGCCCTTCTGTATGTCAATGCCA 540
 DB 481 GAAGGCTCAGTGAAGCAGAACTCTGCGCCAGAAAGCCCTTCTGTATGTCAATGCCA 540
 QY 541 GACCTGATGATCCGGGCACTCCCAATGCGCAGCTTTATTAACAGATTTGATCCAGTT 600
 DB 541 GACCTGATGATCCGGGCACTCCCAATGCGCAGCTTTATTAACAGATTTGATCCAGTT 600
 QY 541 GACCTGATGATCCGGGCACTCCCAATGCGCAGCTTTATTAACAGATTTGATCCAGTT 600
 DB 541 GACCTGATGATCCGGGCACTCCCAATGCGCAGCTTTATTAACAGATTTGATCCAGTT 600
 QY 601 CCATATGATCAACATGTGATGATCTTTCAGATCAACAAACGAGGAGCATCTCTCTT 660

DB 601 CCATATGATCAACATGTGATGATCTTTCAGATCAACAAACGAGGAGCATCTCTCTT 660
 QY 661 ACCCGAGAGGATCTCAGGAATTTGAATCTCTGCTAAGAAATCTTCTATATCTGAGTAC 720
 DB 661 ACCCGAGAGGATCTCAGGAATTTGAATCTCTGCTAAGAAATCTTCTATATCTGAGTAC 720
 QY 721 TCAGTGAAGACATGAGGAGCCAGATGAGAAATTTCTTCAGATATCCACATCTGTGAT 780
 DB 721 TCAGTGAAGACATGAGGAGCCAGATGAGAAATTTCTTCAGATATCCACATCTGTGAT 780
 QY 721 TCAGTGAAGACATGAGGAGCCAGATGAGAAATTTCTTCAGATATCCACATCTGTGAT 780
 DB 721 TCAGTGAAGACATGAGGAGCCAGATGAGAAATTTCTTCAGATATCCACATCTGTGAT 780
 QY 781 ATCATAGTGAACAGAAATTTTGAAGACCAAAACCTGTGAGATGTTGAAAACCTCA 840
 DB 781 ATCATAGTGAACAGAAATTTTGAAGACCAAAACCTGTGAGATGTTGAAAACCTCA 840
 QY 781 ATCATAGTGAACAGAAATTTTGAAGACCAAAACCTGTGAGATGTTGAAAACCTCA 840
 DB 781 ATCATAGTGAACAGAAATTTTGAAGACCAAAACCTGTGAGATGTTGAAAACCTCA 840
 QY 841 ACTGATCTTCACCCCATCAAAATCACTCAGTGGGCTGGAAATATCCCGGTGACAAAT 900
 DB 841 ACTGATCTTCACCCCATCAAAATCACTCAGTGGGCTGGAAATATCCCGGTGACAAAT 900
 QY 841 ACTGATCTTCACCCCATCAAAATCACTCAGTGGGCTGGAAATATCCCGGTGACAAAT 900
 DB 841 ACTGATCTTCACCCCATCAAAATCACTCAGTGGGCTGGAAATATCCCGGTGACAAAT 900
 QY 901 TCCTTAGTTGACAAAGAGAGTGCACAAATCTCCATTTTCAATGACAGAAAGAT 960
 DB 901 TCCTTAGTTGACAAAGAGAGTGCACAAATCTCCATTTTCAATGACAGAAAGAT 960
 QY 901 TCCTTAGTTGACAAAGAGAGTGCACAAATCTCCATTTTCAATGACAGAAAGAT 960
 DB 901 TCCTTAGTTGACAAAGAGAGTGCACAAATCTCCATTTTCAATGACAGAAAGAT 960
 QY 961 ATTTAGGTGACTGAGCCCTTGAACCGAAGAAAGATGATATGTTTATGCAATT 1020
 DB 961 ATTTAGGTGACTGAGCCCTTGAACCGAAGAAAGATGATATGTTTATGCAATT 1020
 QY 961 ATTTAGGTGACTGAGCCCTTGAACCGAAGAAAGATGATATGTTTATGCAATT 1020
 DB 961 ATTTAGGTGACTGAGCCCTTGAACCGAAGAAAGATGATATGTTTATGCAATT 1020
 QY 1021 GCAAGAGATGATGAGGAAACCACTTTCAATTCCTGGAATTCATGTAAGTAAGTTAA 1080
 DB 1021 GCAAGAGATGATGAGGAAACCACTTTCAATTCCTGGAATTCATGTAAGTAAGTTAA 1080
 QY 1021 GCAAGAGATGATGAGGAAACCACTTTCAATTCCTGGAATTCATGTAAGTAAGTTAA 1080
 DB 1021 GCAAGAGATGATGAGGAAACCACTTTCAATTCCTGGAATTCATGTAAGTAAGTTAA 1080
 QY 1081 GATATTATGATATATCACTTACATATGCTGACAGTAACCTATTTAGTCCAGAG 1140
 DB 1081 GATATTATGATATATCACTTACATATGCTGACAGTAACCTATTTAGTCCAGAG 1140
 QY 1081 GATATTATGATATATCACTTACATATGCTGACAGTAACCTATTTAGTCCAGAG 1140
 DB 1081 GATATTATGATATATCACTTACATATGCTGACAGTAACCTATTTAGTCCAGAG 1140
 QY 1141 AATGAAGACTGGGTAACAGATGCGGACCTTTATGCAATGACAGGATGAAGAAAT 1200
 DB 1141 AATGAAGACTGGGTAACAGATGCGGACCTTTATGCAATGACAGGATGAAGAAAT 1200
 QY 1141 AATGAAGACTGGGTAACAGATGCGGACCTTTATGCAATGACAGGATGAAGAAAT 1200
 DB 1141 AATGAAGACTGGGTAACAGATGCGGACCTTTATGCAATGACAGGATGAAGAAAT 1200
 QY 1201 ACTGCCAAGCTTTTCTAACTCAAGATTTGAGACCAATCCCAATCTTCCATGAT 1260
 DB 1201 ACTGCCAAGCTTTTCTAACTCAAGATTTGAGACCAATCCCAATCTTCCATGAT 1260
 QY 1201 ACTGCCAAGCTTTTCTAACTCAAGATTTGAGACCAATCCCAATCTTCCATGAT 1260
 DB 1201 ACTGCCAAGCTTTTCTAACTCAAGATTTGAGACCAATCCCAATCTTCCATGAT 1260
 QY 1261 GGACTCTTCTAATCCAACTATGCTGAAATGTTACAGTTACTTAAACGTCCTGAG 1320
 DB 1261 GGACTCTTCTAATCCAACTATGCTGAAATGTTACAGTTACTTAAACGTCCTGAG 1320
 QY 1261 GGACTCTTCTAATCCAACTATGCTGAAATGTTACAGTTACTTAAACGTCCTGAG 1320
 DB 1261 GGACTCTTCTAATCCAACTATGCTGAAATGTTACAGTTACTTAAACGTCCTGAG 1320
 QY 1321 AAGCAAGATCTCTCAGTACAACTTAAAGATGAGGTCTGACAAAGATTTCAAGACC 1380
 DB 1321 AAGCAAGATCTCTCAGTACAACTTAAAGATGAGGTCTGACAAAGATTTCAAGACC 1380
 QY 1321 AAGCAAGATCTCTCAGTACAACTTAAAGATGAGGTCTGACAAAGATTTCAAGACC 1380
 DB 1321 AAGCAAGATCTCTCAGTACAACTTAAAGATGAGGTCTGACAAAGATTTCAAGACC 1380
 QY 1381 CTTTGTGTTGGAATCAACGTTATGATATCAATGAGATGCCCATCTTGAAGAA 1440
 DB 1381 CTTTGTGTTGGAATCAACGTTATGATATCAATGAGATGCCCATCTTGAAGAA 1440
 QY 1381 CTTTGTGTTGGAATCAACGTTATGATATCAATGAGATGCCCATCTTGAAGAA 1440
 DB 1381 CTTTGTGTTGGAATCAACGTTATGATATCAATGAGATGCCCATCTTGAAGAA 1440
 QY 1441 TCAGATTTAGAAACCTGACTTCTGTAAGACAAACCTTGGGTCCACCATCTTAAACC 1500
 DB 1441 TCAGATTTAGAAACCTGACTTCTGTAAGACAAACCTTGGGTCCACCATCTTAAACC 1500
 QY 1441 TCAGATTTAGAAACCTGACTTCTGTAAGACAAACCTTGGGTCCACCATCTTAAACC 1500
 DB 1441 TCAGATTTAGAAACCTGACTTCTGTAAGACAAACCTTGGGTCCACCATCTTAAACC 1500
 QY 1501 ATCCAGGCACTGATGCTGATGAGCAATTTACTGAGGAGTTCTTAAATTCGTATCATATC 1560
 DB 1501 ATCCAGGCACTGATGCTGATGAGCAATTTACTGAGGAGTTCTTAAATTCGTATCATATC 1560
 QY 1501 ATCCAGGCACTGATGCTGATGAGCAATTTACTGAGGAGTTCTTAAATTCGTATCATATC 1560
 DB 1501 ATCCAGGCACTGATGCTGATGAGCAATTTACTGAGGAGTTCTTAAATTCGTATCATATC 1560
 QY 1561 ATAAAGGAGACAGTGAAGGAGCCCTGAGGGGTGACACAGATCCCATACCAACCGGA 1620
 DB 1561 ATAAAGGAGACAGTGAAGGAGCCCTGAGGGGTGACACAGATCCCATACCAACCGGA 1620
 QY 1561 ATAAAGGAGACAGTGAAGGAGCCCTGAGGGGTGACACAGATCCCATACCAACCGGA 1620
 DB 1561 ATAAAGGAGACAGTGAAGGAGCCCTGAGGGGTGACACAGATCCCATACCAACCGGA 1620
 QY 1621 TATGTCTAATTTAAAGCTCTTGAATTTTGAACAGACAGCTGTTTCCACATTTGTC 1680
 DB 1621 TATGTCTAATTTAAAGCTCTTGAATTTTGAACAGACAGCTGTTTCCACATTTGTC 1680
 QY 1621 TATGTCTAATTTAAAGCTCTTGAATTTTGAACAGACAGCTGTTTCCACATTTGTC 1680
 DB 1621 TATGTCTAATTTAAAGCTCTTGAATTTTGAACAGACAGCTGTTTCCACATTTGTC 1680
 QY 1681 AAAGCAGAAATCTGAGCTTGAAGTGTGTTGATGAGTACATGCAATGCAATCTTGTGCC 1740

Db	1681	AAAGCAGAAATTCCTGAGCCTCTGATGTTTGGTGTGAGTACAAATGCAAGTCTTTTGCC	1741
Qy	1741	AAGTTCACGCTTATTTGTGACAGATGTGAATGAAGACCTCAATTTTCCCAACAGTATTC	1800
Db	1741	AAGTTCACGCTTATTTGTGACAGATGTGAATGAAGACCTCAATTTTCCCAACAGTATTC	1800
Qy	1801	CAAGGGAAGTACAGAGAGATGTACTATAGGCACTAAAGTGGGAAATGTGACTGGCAAG	1866
Db	1801	CAAGGGAAGTACAGAGAGATGTACTATAGGCACTAAAGTGGGAAATGTGACTGGCAAG	1866
Qy	1861	GATCCAGAAGGCTGTGGACATAAGCTATTTCACTGAGGGGAGACACAAGAGTTGGCTTAAA	1920
Db	1861	GATCCAGAAGGCTGTGGACATAAGCTATTTCACTGAGGGGAGACACAAGAGTTGGCTTAAA	1920
Qy	1921	ATTGACACGTCAGTCGTGAGATCTTTATGTGTGCTCCATTTGACAGAGAAAGCCGGAAGT	1980
Db	1921	ATTGACACGTCAGTCGTGAGATCTTTATGTGTGCTCCATTTGACAGAGAAAGCCGGAAGT	1980
Qy	1981	CCATATCGGGATCAAGTGTGGCCACAGAAAGTAAAGGGGGGCTCTTCTTAAAGCTCTGTGTCA	2040
Db	1981	CCATATCGGGATCAAGTGTGGCCACAGAAAGTAAAGGGGGGCTCTTCTTAAAGCTCTGTGTCA	2040
Qy	2041	GAGTTCACACCTGATCTTTATGTGATGTGAATGCAACCTCCAGGCTGACCAAGGACTAC	2100
Db	2041	GAGTTCACACCTGATCTTTATGTGATGTGAATGCAACCTCCAGGCTGACCAAGGACTAC	2100
Qy	2101	ACGGGCTTGTTCTTCTGCCATCCCTCAGTGCACCTGGAAGTCTCATTTTGAGGCTACT	2160
Db	2101	ACGGGCTTGTTCTTCTGCCATCCCTCAGTGCACCTGGAAGTCTCATTTTGAGGCTACT	2160
Qy	2161	GATGATGATCAGCATTTATTTGGGGTCCCATTTTACATTTTCCCTGGCAGTGAAGC	2220
Db	2161	GATGATGATCAGCATTTATTTGGGGTCCCATTTTACATTTTCCCTGGCAGTGAAGC	2220
Qy	2221	TTACAAAACGACTGGGAAATTTCAAATAATGATGCTCAATGGCCGACCTGTACAGAG	2280
Db	2221	TTACAAAACGACTGGGAAATTTCAAATAATGATGCTCAATGGCCGACCTGTACAGAG	2280
Qy	2281	CACACAGACTTTGAGAGAGGGCGGTATGCTGTCTGATCCGCATCAATGATGAGGGGGTGG	2340
Db	2281	CACACAGACTTTGAGAGAGGGCGGTATGCTGTCTGATCCGCATCAATGATGAGGGGGTGG	2340
Qy	2341	CCACCCCTTGGAGGCAATGTTCTTTTACAGATTCATCTCAGTGTGTGTGAAGGAAGT	2400
Db	2341	CCACCCCTTGGAGGCAATGTTCTTTTACAGATTCATCTCAGTGTGTGTGAAGGAAGT	2400
Qy	2401	TGTTTCGGGGCAGCAGGTCAACAGACTGGGAAATCCACTGTGGGAGATGGCAGTTGGTATA	2460
Db	2401	TGTTTCGGGGCAGCAGGTCAACAGACTGGGAAATCCACTGTGGGAGATGGCAGTTGGTATA	2460
Qy	2461	CTGCTGACCAACCCTCTGTGTGATTTGGTATTAATTTTATGACGTTGTGTTATCCGATAAAG	2520
Db	2461	CTGCTGACCAACCCTCTGTGTGATTTGGTATTAATTTTATGACGTTGTGTTATCCGATAAAG	2520
Qy	2521	AAGGATAAAGGCAAGATPATGTGAAATGTGTAAGTGTCTCAAGATCTGAAGTCAAACTCTGAGA	2580
Db	2521	AAGGATAAAGGCAAGATPATGTGAAATGTGTAAGTGTCTCAAGATCTGAAGTCAAACTCTGAGA	2580
Qy	2581	AGCTGAATTTGAAAAGGAATGTTTGAATTTATATGCAAGTGTCTATTTACAGAACCA	2640
Db	2581	AGCTGAATTTGAAAAGGAATGTTTGAATTTATATGCAAGTGTCTATTTACAGAACCA	2640
Qy	2641	TCTCATCTTATTAATCTTTTCACTTAAAGTCAATTAATTTTATTTTAAACAGATATCCCTCT	2700
Db	2641	TCTCATCTTATTAATCTTTTCACTTAAAGTCAATTAATTTTATTTTAAACAGATATCCCTCT	2700
Qy	2701	TGTCCTTAAATATTTGCTAAATATTTCTTTTATGAGTGAAGTCTGTCTGTGCCAG	2760
Db	2701	TGTCCTTAAATATTTGCTAAATATTTCTTTTATGAGTGAAGTCTGTCTGTGCCAG	2760
Qy	2761	GCTGAGATACAGTGTGTGATTCACAGCTCACTGCAACCTCCGCTCTCCGGGTTACATGA	2820
Db	2761	GCTGAGATACAGTGTGTGATTCACAGCTCACTGCAACCTCCGCTCTCCGGGTTACATGA	2820

QY	2821	TTCTCCGCGCCAGGCTTCCTTAAGTACGTGGGTTTACAGGCAACCCACACATGCCAGCT	2880		
Db	2821	TTCTCTGCGCTCAGGCTCTCTAAGTACGTGGGTTTACAGGCAACCCACACATGCCAGCT	2880		
QY	2881	AATTTTGTATTTTAAATAGAGACGGGGTTTCCGCATTTGGCCAGGCTGGTCTTGAATC	2940		
Db	2881	AATTTTGTATTTTAAATAGAGACGGGGTTTCCGCATTTGGCCAGGCTGGTCTTGAATC	2940		
QY	2941	CTGACGTCAATGATCTGCGCTTGGTCTCCCAATACAGGCAATGAACCATGCAACCA	3000		
Db	2941	CTGACGTCAATGATCTGCGCTTGGTCTCCCAATACAGGCAATGAACCATGCAACCA	3000		
QY	3001	CCCTACTAGATATTTCATGTCCTATAGACATATAGAGAGATTTTTCATTTTCCATGACAT	3060		
Db	3001	CCCTACTAGATATTTCATGTCCTATAGACATATAGAGAGATTTTTCATTTTCCATGACAT	3060		
QY	3061	TTTTCTCTCGCAAAATGGCTTAGCTACCTTGTTTTCCTTTTGGGGCAAGCAGACT	3120		
Db	3061	TTTTCTCTCTCGCAAAATGGCTTAGCTACCTTGTTTTCCTTTTGGGGCAAGCAGACT	3120		
QY	3121	CATTAAATATTCTGTACATTTTCTTTATCAGAGATATATCATGTTGTCTCATAGA	3180		
Db	3121	CATTAAATATTCTGTACATTTTCTTTATCAGAGATATATCATGTTGTCTCATAGA	3180		
QY	3181	ACTGCGCGGATTCGATTTATGTTTTTCTGATTCATCTCGTGGCCCTCATCTTGAC	3240		
Db	3181	ACTGCGCGGATTCGATTTATGTTTTTCTGATTCATCTCGTGGCCCTCATCTTGAC	3240		
QY	3241	TCCTTGGTATTTTCAGCAATTTCCAAACATTTGTCCAGAGAAAGAAAGTGAGACTCAG	3300		
Db	3241	TCCTTGGTATTTTCAGCAATTTCCAAACATTTGTCCAGAGAAAGAAAGTGAGACTCAG	3300		
QY	3301	GAAGAAATTAATTAATAAAGACAGCCTTTTGGGCGCGCAATTC	3345		
Db	3301	GAAGAAATTAATTAATAAAGACAGCCTTTTGGGCGCGCAATTC	3345		
RESULT 4					
ABL62444	standard; DNA, 3345 BP.				
XX	ABL62444;				
AC					
XX					
DT	15-MAY-2002	(first entry)			
DE	Colon adenocarcinoma related gene sequence SEQ ID NO: 781.				
XX					
KM	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;				
KW	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;				
KM	cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;				
KW	gene; ds.				
XX					
OS	Homo sapiens.				
XX					
PN	MO2001.94629-A2.				
XX					
PD	13-DEC-2001.				
XX					
PF	30-MAY-2001;	2001WO-US10838.			
XX					
PR	05-JUN-2000;	2000US-2095473P.			
PR	05-JUN-2000;	2000US-2095473P.			
PR	18-SEP-2000;	2000US-233133P.			
PR	18-SEP-2000;	2000US-233133P.			
PR	20-SEP-2000;	2000US-234034P.			
PR	20-SEP-2000;	2000US-234034P.			
PR	22-SEP-2000;	2000US-234509P.			
PR	22-SEP-2000;	2000US-234509P.			
PR	23-SEP-2000;	2000US-234923P.			
PR	23-SEP-2000;	2000US-234923P.			
PR	25-SEP-2000;	2000US-235077P.			
PR	25-SEP-2000;	2000US-235077P.			

PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 PA (AVAL-) AVALON PHARM.
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppe DR, Weaver Z;
 XX WPI; 2002-188264/24.
 XX
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 PS Claim 1; SEQ ID 781; 44pp; English.
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (1) of a signature gene set, where (1)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL6164
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (1) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 XX
 XX Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;
 SQ
 Query Match 100.0%; Score 3345; DB 24; Length 3345;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 61 CAGAAAGAAATTGAAATAAGAAAAGAACTATGATCTTCAGGCCATCTTCACCTCTGTGT 120
 QY 121 CTTCCTATGCTTTATTTGGCACTGATATGCGCAGAGGGAGATTGTGACCCCTG 180
 DB 121 CTTCCTATGCTTTATTTGGCACTGATATGCGCAGAGGGAGATTGTGACCCCTG 180
 QY 181 AAGCCATGACATTTTCTATTTATGAAAGGCAAGGCAAGTCAATTAATTCAGATT 240
 DB 181 AAGCCATGACATTTTCTATTTATGAAAGGCAAGGCAAGTCAATTAATTCAGATT 240
 QY 241 AAGCCATGCTCTCTGCTGTGACTTTTGAATTAAGGGGAGACAGCAATATTTGTG 300
 DB 241 AAGCCATGCTCTCTGCTGTGACTTTTGAATTAAGGGGAGACAGCAATATTTGTG 300
 QY 301 ATGAAACGGAGGAGACTTCTGTATTAACAAGAGCCTTGGACAGGGAACAAGATCTACT 360
 DB 301 ATGAAACGGAGGAGACTTCTGTATTAACAAGAGCCTTGGACAGGGAACAAGATCTACT 360
 QY 361 CACAATCTCCAGGTTGAGCCCTGAGCGCTAATGAAATTAATGAGGGTCCAGTCCCT 420
 DB 361 CACAATCTCCAGGTTGAGCCCTGAGCGCTAATGAAATTAATGAGGGTCCAGTCCCT 420
 QY 421 ATCACCATAGAAAGTGAAGGACATCAACGACATCGACCGTTTCTCCAGTCAAGTAC 480
 DB 421 ATCACCATAGAAAGTGAAGGACATCAACGACATCGACCGTTTCTCCAGTCAAGTAC 480
 QY 481 GAAGGCTCAGTAAAGGAGAACTCTGCGCCAGGAAAGCCCTCTTGTATGTCATGCGACA 540
 DB 481 GAAGGCTCAGTAAAGGAGAACTCTGCGCCAGGAAAGCCCTCTTGTATGTCATGCGACA 540
 QY 541 GACCTGATGATCCGGGCACTCCCAATGGCAGCTTTATTAACGATTTGTATCATCAGCTT 600
 DB 541 GACCTGATGATCCGGGCACTCCCAATGGCAGCTTTATTAACGATTTGTATCATCAGCTT 600
 QY 601 CCCATGATCAACAATGTCATGTAATTTGAGATCAACAACAAACGGAGGCAATCTCTCT 660
 DB 601 CCCATGATCAACAATGTCATGTAATTTGAGATCAACAACAAACGGAGGCAATCTCTCT 660
 QY 661 ACCCGAGAGGAGTCTCAGAAATTTGAATTCCTGCTAAGAAATCCCTATATCTGAGTAC 720
 DB 661 ACCCGAGAGGAGTCTCAGAAATTTGAATTCCTGCTAAGAAATCCCTATATCTGAGTAC 720
 QY 721 TCAAGTGAAGGACATGGAGGCGCAGAGTGAATTCCTTCAGTATACCATCTGTGAT 780
 DB 721 TCAAGTGAAGGACATGGAGGCGCAGAGTGAATTCCTTCAGTATACCATCTGTGAT 780
 QY 781 ATCATAGTGAACAGAAATATTTGGAAAGCAACAAACCTGTGAGATGTGGAACCTCA 840
 DB 781 ATCATAGTGAACAGAAATATTTGGAAAGCAACAAACCTGTGAGATGTGGAACCTCA 840
 QY 841 ACTGATCTCAACCCCATCAAAATCACTCAGGTGGGTGGAATGATCCCGGTGCAATAT 900
 DB 841 ACTGATCTCAACCCCATCAAAATCACTCAGGTGGGTGGAATGATCCCGGTGCAATAT 900
 QY 901 TCCCTAGTGAACAAAGAGAGTGCAGAGATTCCTCATTTTCAATTGACAGGAGAT 960
 DB 901 TCCCTAGTGAACAAAGAGAGTGCAGAGATTCCTCATTTTCAATTGACAGGAGAT 960
 QY 961 ATTTAGTGAATCAGCCCTTGAACCGAAGAAAGATGATGATGTTTTTATGACGTT 1020
 DB 961 ATTTAGTGAATCAGCCCTTGAACCGAAGAAAGATGATGATGTTTTTATGACGTT 1020
 QY 1021 GCAAAAGATGATGAGGAAACCACTTTTCAATTCCTGGAATTCATGTAAGTAAAGTAAA 1080
 DB 1021 GCAAAAGATGATGAGGAAACCACTTTTCAATTCCTGGAATTCATGTAAGTAAAGTAAA 1080
 QY 1081 GATATTAATGATTAATCACTTACATGTCCTGACCAATACGATTTGAGGTCCAGAG 1140
 DB 1081 GATATTAATGATTAATCACTTACATGTCCTGACCAATACGATTTGAGGTCCAGAG 1140
 QY 1141 AATGAGACTGGGTAAACGATATGAGGACCTTATCTGACATGACAGGATGAAGAAAT 1200
 DB 1141 AATGAGACTGGGTAAACGATATGAGGACCTTATCTGACATGACAGGATGAAGAAAT 1200

Db 1141 AATGAAGACTGGGTAACGATATCGGACCTTACTGCACATGAACGGGATGAAGAAAT 1200
Qy 1201 ACTGCACAAAGTTTCTTAAACTACAGATGTGGAGAAACTCCCAACTTCCCATGGAT 1260
Db 1201 ACTGCCAAAGTTTCTTAAACTACAGATGTGGAGAAACTCCCAACTTCCCATGGAT 1260
Qy 1261 GGACTCTTCTTAATCCAACTATGTGGAAATGTTCAGATTAGCTTAAACAGTCTTGAAG 1320
Db 1261 GGACTCTTCTTAATCCAACTATGTGGAAATGTTCAGATTAGCTTAAACAGTCTTGAAG 1320
Qy 1321 AAGCAATATCTCTCGTACAACTTAAACGATAGAGGTCTGACAAAGATTTCAGAAC 1380
Db 1321 AAGCAATATCTCTCGTACAACTTAAACGATAGAGGTCTGACAAAGATTTCAGAAC 1380
Qy 1381 CTTGTGTTTGTGCAAACTCAAGTTATGATATCAATGATCAGATCCCATCTTTGAAAA 1440
Db 1381 CTTGTGTTTGTGCAAACTCAAGTTATGATATCAATGATCAGATCCCATCTTTGAAAA 1440
Qy 1441 TCGATATATGAAAACCTGACTCTTGTGTAAGACAAACATTTGGGTCCACCATCTTAAC 1500
Db 1441 TCGATATATGAAAACCTGACTCTTGTGTAAGACAAACATTTGGGTCCACCATCTTAAC 1500
Qy 1501 ATCCAGGCCACTGATGTGCTGATAGGCCATTTACTGGGAGTTCTAAATTCGTATCATATC 1560
Db 1501 ATCCAGGCCACTGATGTGCTGATAGGCCATTTACTGGGAGTTCTAAATTCGTATCATATC 1560
Qy 1561 ATAAAGGAGACAGTAGAGGACGCGCTGGGGGTTGACACAGATCCCATACCAACCGGA 1620
Db 1561 ATAAAGGAGACAGTAGAGGACGCGCTGGGGGTTGACACAGATCCCATACCAACCGGA 1620
Qy 1621 TATGTCATTAATVAAAAAGCTCTTGATTTTGAACAGACAGCTGTTTCCAACTTGTGTC 1680
Db 1621 TATGTCATTAATVAAAAAGCTCTTGATTTTGAACAGACAGCTGTTTCCAACTTGTGTC 1680
Qy 1681 AAAGCAGAAAATCCTGAGCCTCTAGTGTGTCGTGTAAGTACATGCAATGCTTTTGGC 1740
Db 1681 AAAGCAGAAAATCCTGAGCCTCTAGTGTGTCGTGTAAGTACATGCAATGCTTTTGGC 1740
Qy 1741 AAGTTCACGCTTATTTGACAGATGTGATGATGAAGACCTCAATTTTCCCAACGATATTC 1800
Db 1741 AAGTTCACGCTTATTTGACAGATGTGATGATGAAGACCTCAATTTTCCCAACGATATTC 1800
Qy 1801 CAAGCAGAAAGTCAAGTAGAGATGTAGCTAAGGACCTAAGTGGGCAATGTACTGCCAG 1860
Db 1801 CAAGCAGAAAGTCAAGTAGAGATGTAGCTAAGGACCTAAGTGGGCAATGTACTGCCAG 1860
Qy 1861 GATCCAGAAAGTCTGACATATAGCTATTCCTGAGGGGAGACACAGAGGTGGCTTAA 1920
Db 1861 GATCCAGAAAGTCTGACATATAGCTATTCCTGAGGGGAGACACAGAGGTGGCTTAA 1920
Qy 1921 ATTGACCAAGTGAAGTGTAGATCTTATGTGGGCTCATTTGACAGAGACCGGAAGT 1980
Db 1921 ATTGACCAAGTGAAGTGTAGATCTTATGTGGGCTCATTTGACAGAGACCGGAAGT 1980
Qy 1981 CCATATCGGGGTACAAAGTGTGGCCACAGAAAGTGGGGGCTCTCTTAAGCTGTGTCA 2040
Db 1981 CCATATCGGGGTACAAAGTGTGGCCACAGAAAGTGGGGGCTCTCTTAAGCTGTGTCA 2040
Qy 2041 GACTTTCACCTGATCTTATGATGTGATGATGACAACTCTCCAGGCTAGCAAGACTAC 2100
Db 2041 GACTTTCACCTGATCTTATGATGTGATGATGACAACTCTCCAGGCTAGCAAGACTAC 2100
Qy 2101 ACGGGCTTGTCTTCTGCGCACTCCCTCAGTGCACCTGGAAGTCTCATTTTCAGAGCTACT 2160
Db 2101 ACGGGCTTGTCTTCTGCGCACTCCCTCAGTGCACCTGGAAGTCTCATTTTCAGAGCTACT 2160
Qy 2161 GATGATGATCAGACATTAATTTGGGGGTCCCATTTTACATTTTCCCTCGGCACTGAGAGC 2220
Db 2161 GATGATGATCAGACATTAATTTGGGGGTCCCATTTTACATTTTCCCTCGGCACTGAGAGC 2220
Qy 2221 TTACAAAAGCAGCTGGGAAGTTTCCAAATCAATGTGATCTCATGCTCCGAGCTGTACAGG 2280
Db 2221 TTACAAAAGCAGCTGGGAAGTTTCCAAATCAATGTGATCTCATGCTCCGAGCTGTACAGG 2280

Qy 2281 CACACAGACTTTGAGAGAGAGGGCGTATGTGCTTGAATCCGATCAATGATGGGGGTGCG 2340
Db 2281 CACACAGACTTTGAGAGAGAGGGCGTATGTGCTTGAATCCGATCAATGATGGGGGTGCG 2340
Qy 2341 CCAACCTTGGAGAGCATTGTTCTTTTCCAGATTACTTGCAGTGTGTGGAAGAGT 2400
Db 2341 CCAACCTTGGAGAGCATTGTTCTTTTCCAGATTACTTGCAGTGTGTGGAAGAGT 2400
Qy 2401 TGTTCGCGGACAGAGTCAACGACTGGGATACCACTGTGGCATGTGAGATTGGTATTA 2460
Db 2401 TGTTCGCGGACAGAGTCAACGACTGGGATACCACTGTGGCATGTGAGATTGGTATTA 2460
Qy 2461 CTGCTGACCAACCTTCTGCTGATATTTGATTAATTTAGCAGTGTGTTATCCGATTAAG 2520
Db 2461 CTGCTGACCAACCTTCTGCTGATATTTGATTAATTTAGCAGTGTGTTATCCGATTAAG 2520
Qy 2521 AAGGATPAAGGCAAGATTAATTTGAAAGTCTCAAGCATCTGAAGTCAAACTCTGAGA 2580
Db 2521 AAGGATPAAGGCAAGATTAATTTGAAAGTCTCAAGCATCTGAAGTCAAACTCTGAGA 2580
Qy 2581 AGCTGAATTTGAAAAGGAATGTTGAATTTATATAGCAAGTCTATTTTCAGCAACCA 2640
Db 2581 AGCTGAATTTGAAAAGGAATGTTGAATTTATATAGCAAGTCTATTTTCAGCAACCA 2640
Qy 2641 TCTCATCTATTAATTTTCTTCTCACTAAGTGTGATTAATTTTAAACAGATATTCCTCT 2700
Db 2641 TCTCATCTATTAATTTTCTTCTCACTAAGTGTGATTAATTTTAAACAGATATTCCTCT 2700
Qy 2701 TGTCTTTAATTAATTTGCTAAATATTTCTTTTGAAGTGGAGTCTGCTGTGCGCCAG 2760
Db 2701 TGTCTTTAATTAATTTGCTAAATATTTCTTTTGAAGTGGAGTCTGCTGTGCGCCAG 2760
Qy 2761 GCTGAGTACAGTGTGTGATATCCAGCTCACTGCAACCTCCGCTCTGAGTTCATGTA 2820
Db 2761 GCTGAGTACAGTGTGTGATATCCAGCTCACTGCAACCTCCGCTCTGAGTTCATGTA 2820
Qy 2821 TTCTCTGCTCAGCTTCTCAAGTATGCTGGGTTTACAGGACCCACACATGCGCAGCT 2880
Db 2821 TTCTCTGCTCAGCTTCTCAAGTATGCTGGGTTTACAGGACCCACACATGCGCAGCT 2880
Qy 2881 AATTTTGTATTTTAAATAGACAGGGGTTTGGCCATTTGGCCAGGCTGTGTAATTC 2940
Db 2881 AATTTTGTATTTTAAATAGACAGGGGTTTGGCCATTTGGCCAGGCTGTGTAATTC 2940
Qy 2941 CTGAGCTCAAGTATCTGCTGCTGTGCTCTCCAAATACAGCATTAACCACTGCACCA 3000
Db 2941 CTGAGCTCAAGTATCTGCTGCTGTGCTCTCCAAATACAGCATTAACCACTGCACCA 3000
Qy 3001 CCTACTTAATTTTCAATGAGTATGATGATTAAGAGATTTTTCATTTTCCATGACAT 3060
Db 3001 CCTACTTAATTTTCAATGAGTATGATGATTAAGAGATTTTTCATTTTCCATGACAT 3060
Qy 3061 TTTTCTCTCTGCAATGAGCTTAGCTATGTGTTTCTTCTTTGGGCAAGACAGACT 3120
Db 3061 TTTTCTCTCTGCAATGAGCTTAGCTATGTGTTTCTTCTTTGGGCAAGACAGACT 3120
Qy 3121 CATTAATATTTCTGATCATTTTCTTTATCAAGAGATATATCAGTGTGTCTCATAGA 3180
Db 3121 CATTAATATTTCTGATCATTTTCTTTATCAAGAGATATATCAGTGTGTCTCATAGA 3180
Qy 3181 ACTGCTGATTCATTAATTTTCTTCTGATTCATCTGTGTCTTCAATCCCTGAGC 3240
Db 3181 ACTGCTGATTCATTAATTTTCTTCTGATTCATCTGTGTCTTCAATCCCTGAGC 3240
Qy 3241 TCTTTGTATTTTCACTGAATTTTCAAACTTTGTCAAGAGAAAATAATGAGACTCAG 3300
Db 3241 TCTTTGTATTTTCACTGAATTTTCAAACTTTGTCAAGAGAAAATAATGAGACTCAG 3300
Qy 3301 GAAAAATTAATTAATTAAGAACAGCTTTTGGGCGCGGAATTC 3345
Db 3301 GAAAAATTAATTAATTAAGAACAGCTTTTGGGCGCGGAATTC 3345

RESULT 5
ID ABL64424
PT ABL64424 standard; DNA; 3345 BP.
XX
AC ABL64424;
XX
DT 15-MAY-2002 (first entry)
XX
DE Stomach cancer related gene sequence SEQ ID NO:2761.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW Stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW Cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN W0200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209472P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-231133P.
PR 18-SEP-2000; 2000US-231617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235083P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Sopet DR, Weaver Z;
XX
DE WPI; 2002-188264/24.

XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
XX
PS Claim 1; SEQ ID 2761; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL64
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;
Query Match 100.0%; Score 3345; DB 24; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAATTCGGTCTGACCATGATGAGAAAGACCTTTAAACCATTTGTGACTTA 60
DB 1 GAATTCGGTCTGACCATGATGAGAAAGACCTTTAAACCATTTGTGACTTA 60
QY 61 CAGAAAGAAATTTGAATTAAGAAATATGATATCTTCAGGCCATCTTCACTCCCTGTG 120
DB 61 CAGAAAGAAATTTGAATTAAGAAATATGATATCTTCAGGCCATCTTCACTCCCTGTG 120
QY 121 CTTCCTATGCTTTATTTGGCACTGATGAGCAAGAGGGGAAAGTTAGTGACCCCTG 180
DB 121 CTTCCTATGCTTTATTTGGCACTGATGAGCAAGAGGGGAAAGTTAGTGACCCCTG 180
QY 181 AAACCATGACATTTCTATTATTTAGAGGCAAGACCGAGTCAATATATATCCAGTT 240
DB 181 AAACCATGACATTTCTATTATTTAGAGGCAAGACCGAGTCAATATATATCCAGTT 240
QY 241 AAGGCCAATCTCTGCTGTGACTTTGAACTAAGTGGGAGACAGCAACATATTTGTG 300
DB 241 AAGGCCAATCTCTGCTGTGACTTTGAACTAAGTGGGAGACAGCAACATATTTGTG 300
QY 301 ATGAAACGGAGGAGACTTCTGTATTCACAGAGCCTTGGACAGGAAACAAATTTACT 360
DB 301 ATGAAACGGAGGAGACTTCTGTATTCACAGAGCCTTGGACAGGAAACAAATTTACT 360
QY 361 CACAATTCAGAGTTGAGCCCTGAGCGGTAATGAAATTAATGAGAGGCTCCAGTCCCT 420
DB 361 CACAATTCAGAGTTGAGCCCTGAGCGGTAATGAAATTAATGAGAGGCTCCAGTCCCT 420
QY 421 ATCACCATTAGAAGTGAAGACATCAACGACATGACCCAGTTTCTCCAGTCAAGTAC 480
DB 421 ATCACCATTAGAAGTGAAGACATCAACGACATGACCCAGTTTCTCCAGTCAAGTAC 480
QY 481 GAAGGCTCAGTAAAGCAAACTCTCGCCCGAGAAAGCCCTTCTGTATGTCATGCCACA 540
DB 481 GAAGGCTCAGTAAAGCAAACTCTCGCCCGAGAAAGCCCTTCTGTATGTCATGCCACA 540
QY 541 GACCTGATATATCCGACATCCCAATGAGCGAGTTTATTAACGATTTGATCAGCTT 600
DB 541 GACCTGATATATCCGACATCCCAATGAGCGAGTTTATTAACGATTTGATCAGCTT 600
QY 601 CCCATGATCAACATGATCATGTAATTTAGATCAACAAACGAGGAGCATCTCTT 660
DB 601 CCCATGATCAACATGATCATGTAATTTAGATCAACAAACGAGGAGCATCTCTT 660

D6 601 CCCATGATCAACAAATGTCATCTTCTTCAATCAACAAACAAACGGAGCCATCTCTT 660
QY 661 ACCGAGAGGGATCTCAGAAATGAATCCTCTAAGAAATCCTTCTTAATCTGTATC 720
D6 661 ACCGAGAGGGATCTCAGAAATGAATCCTCTAAGAAATCCTTCTTAATCTGTATC 720
QY 721 TCAGTGAAGACATGGAGGCGCAGAGTGAATTCCTTCACTGATACCAATCTGTGAT 780
D6 721 TCAGTGAAGACATGGAGGCGCAGAGTGAATTCCTTCACTGATACCAATCTGTGAT 780
QY 781 ATCATAGTGAACAGAAATATTTGGAAGACCAAAACCTGTGAGATGTGTGAAGAACTCA 840
D6 781 ATCATAGTGAACAGAAATATTTGGAAGACCAAAACCTGTGAGATGTGTGAAGAACTCA 840
QY 841 ACTGATCCTCAACCCATCAAAATCACTCAGGTGCGGTGAATGATCCCGGTGCACAATAT 900
D6 841 ACTGATCCTCAACCCATCAAAATCACTCAGGTGCGGTGAATGATCCCGGTGCACAATAT 900
QY 901 TCCTTAGTTGACAAAGAGAGCTGCCAAGATTCCTCATTTTCAATTTGACAGAGAGAT 960
D6 901 TCCTTAGTTGACAAAGAGAGCTGCCAAGATTCCTCATTTTCAATTTGACAGAGAGAT 960
QY 961 ATTACGTGACTGACCCCTTGGACCGAGAGAAAGAAAGATGATGTTTTTATGCACTT 1020
D6 961 ATTACGTGACTGACCCCTTGGACCGAGAGAAAGAAAGATGATGTTTTTATGCACTT 1020
QY 1021 GCAAAAGATGATGCGAAAAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTAA 1080
D6 1021 GCAAAAGATGATGCGAAAAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTAA 1080
QY 1081 GATATTAATGATTAATCCACCTACATGTCCTGACACAGTAACCGTATTTGAGGTCCAGAG 1140
D6 1081 GATATTAATGATTAATCCACCTACATGTCCTGACACAGTAACCGTATTTGAGGTCCAGAG 1140
QY 1141 AATGAAGACTGGGTAAACATGTGAGGACCCCTTACTGCAATGACAGAGATGAAGAAAT 1200
D6 1141 AATGAAGACTGGGTAAACATGTGAGGACCCCTTACTGCAATGACAGAGATGAAGAAAT 1200
QY 1201 ACTGCCAAGATTTTCTAAACTACAGAGATTTGGAGCAAACTCCCAAACTTCCATGAT 1260
D6 1201 ACTGCCAAGATTTTCTAAACTACAGAGATTTGGAGCAAACTCCCAAACTTCCATGAT 1260
QY 1261 GGACTCTTCTTAATCCAAACCTATGCTGGAATGTTACAGTTAGCTAAACGTCTTGAAG 1320
D6 1261 GGACTCTTCTTAATCCAAACCTATGCTGGAATGTTACAGTTAGCTAAACGTCTTGAAG 1320
QY 1321 AAGCAAGATATCTCTCAGTACAACTTAAACGATAGAGGTGTGCAAAAGATTTCAAGACC 1380
D6 1321 AAGCAAGATATCTCTCAGTACAACTTAAACGATAGAGGTGTGCAAAAGATTTCAAGACC 1380
QY 1381 CTTGTTTGTGCAAAATCAAGTATATGATATCAATGATCAGATCCCATCTTTGAAAAA 1440
D6 1381 CTTGTTTGTGCAAAATCAAGTATATGATATCAATGATCAGATCCCATCTTTGAAAAA 1440
QY 1441 TCAGATTTATGAAAACCTGACTCTTGTGAAGACAAACCAATGGGTCCACCATCTTAAAC 1500
D6 1441 TCAGATTTATGAAAACCTGACTCTTGTGAAGACAAACCAATGGGTCCACCATCTTAAAC 1500
QY 1501 ATCCAGGCCATGATGCTGATGAGCCATTTACTGGAGATTTCTAAATTTCTGATCATATC 1560
D6 1501 ATCCAGGCCATGATGCTGATGAGCCATTTACTGGAGATTTCTAAATTTCTGATCATATC 1560
QY 1561 ATAAAAGGAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCATACCAACCGGGA 1620
D6 1561 ATAAAAGGAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCATACCAACCGGGA 1620
QY 1621 TATGTATTAATTAATAAGCTCTTGAATTTTGAACAGACAGCTGTTTCAACATTTGTGTC 1680
D6 1621 TATGTATTAATTAATAAGCTCTTGAATTTTGAACAGACAGCTGTTTCAACATTTGTGTC 1680
QY 1681 AAGGAGAAAATCTGAGCTCTAGAGTTTGTGTGTAAGTACATGCAATGTTCTTTTGC 1740
D6 1681 AAGGAGAAAATCTGAGCTCTAGAGTTTGTGTGTAAGTACATGCAATGTTCTTTTGC 1740
QY 1741 AAGTTCAGCTTATTTGACAGATGTAATGAAGACCTCAATTTTCCCAACAGATATC 1800
D6 1741 AAGTTCAGCTTATTTGACAGATGTAATGAAGACCTCAATTTTCCCAACAGATATC 1800
QY 1801 CAAGCGAAGTCAATGAGATGATGATTAAGGACCTAAAGTGGGCAATGTGACTGCCAAG 1860
D6 1801 CAAGCGAAGTCAATGAGATGATGATTAAGGACCTAAAGTGGGCAATGTGACTGCCAAG 1860
QY 1861 GATCCGAAGTCTGACATTAAGCTATTCTGAGGGGAGACCAAGAGTTGGCTTAA 1920
D6 1861 GATCCGAAGTCTGACATTAAGCTATTCTGAGGGGAGACCAAGAGTTGGCTTAA 1920
QY 1921 ATTGACCAAGTCAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
D6 1921 ATTGACCAAGTCAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
QY 1981 CCATATCGGGTACAGTGGTGGCCACAGAAATGAGGGGGGCTTCTTTAAGCTGTGTCA 2040
D6 1981 CCATATCGGGTACAGTGGTGGCCACAGAAATGAGGGGGGCTTCTTTAAGCTGTGTCA 2040
QY 2041 GAGTTCACCTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
D6 2041 GAGTTCACCTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
QY 2101 ACGGGCTTGTCTTCTGCCATCCCTCAGTGCACCTGGAAGTCTCATTTTGGAGGCTACT 2160
D6 2101 ACGGGCTTGTCTTCTGCCATCCCTCAGTGCACCTGGAAGTCTCATTTTGGAGGCTACT 2160
QY 2161 GATGATGATCAGCACTTATTTGCGGGGTCCCATTTTACATTTTCCCTGCGCAGTGAAGC 2220
D6 2161 GATGATGATCAGCACTTATTTGCGGGGTCCCATTTTACATTTTCCCTGCGCAGTGAAGC 2220
QY 2221 TTAACAAACGACCTGGGAAGTTTCCAAATCAATGATGATGATGATGATGATGATGATGAT 2280
D6 2221 TTAACAAACGACCTGGGAAGTTTCCAAATCAATGATGATGATGATGATGATGATGATGAT 2280
QY 2281 CACACAGACTTGAAGAGAGGGGTGATGCTGATGATGATGATGATGATGATGATGATGATGAT 2340
D6 2281 CACACAGACTTGAAGAGAGGGGTGATGCTGATGATGATGATGATGATGATGATGATGATGAT 2340
QY 2341 CCAACCTTGAAGAGCAATGTTCTTTTACAGTATCAATTCGACAGTGTGTGAAGAGAT 2400
D6 2341 CCAACCTTGAAGAGCAATGTTCTTTTACAGTATCAATTCGACAGTGTGTGAAGAGAT 2400
QY 2401 TGTTCCTGGCCAGACAGTCAACAGACTGGGATACCACTGTGGGCAATGGCAGTTGTAT 2460
D6 2401 TGTTCCTGGCCAGACAGTCAACAGACTGGGATACCACTGTGGGCAATGGCAGTTGTAT 2460
QY 2461 CTGCTGACCAACCTTCTGGGATGATGATTAATTTTGAAGTGTGTATTCGCAATTAAG 2520
D6 2461 CTGCTGACCAACCTTCTGGGATGATGATTAATTTTGAAGTGTGTATTCGCAATTAAG 2520
QY 2521 AAGGATTAAGGCAAGATATGTTGAAGGTCTCAAGCATCTGAAGTCAAACTCTGAGA 2580
D6 2521 AAGGATTAAGGCAAGATATGTTGAAGGTCTCAAGCATCTGAAGTCAAACTCTGAGA 2580
QY 2581 AGCTGAATTTGAAAAGAAATGTTGAATTTATATAGCAAGTGTATTTGACAAACCA 2640
D6 2581 AGCTGAATTTGAAAAGAAATGTTGAATTTATATAGCAAGTGTATTTGACAAACCA 2640
QY 2641 TCTCATCTATTAATCTTTCATGATCAAGTCAATTAATTTTAAACATATTCCTCTC 2700
D6 2641 TCTCATCTATTAATCTTTCATGATCAAGTCAATTAATTTTAAACATATTCCTCTC 2700
QY 2701 TGTCTTTAATATTTGCTAAATATTTCTTTTGAAGTGAAGTGTGCTGTGCTGCCAG 2760
D6 2701 TGTCTTTAATATTTGCTAAATATTTCTTTTGAAGTGAAGTGTGCTGTGCTGCCAG 2760
QY 2761 GCTGAGATCAAGTGTGTATCCAGCTCACTGCAACCTTCCGCTCTGAGTTCAATGA 2820
D6 2761 GCTGAGATCAAGTGTGTATCCAGCTCACTGCAACCTTCCGCTCTGAGTTCAATGA 2820

QY	2821	TTCTCCGCGCCACAGTTCTCTAAGTAGCTGGGGTTTACAGAGCACCCACACAGTCCCAAGT	2880
Db	2821	TTCTCTCGCCCTCAGCTCTCTAAGTAGCTGGGGTTTACAGAGCACCCACACAGTCCCAAGT	2880
QY	2881	AATTTTGTATTTTAAATATAGAGACGGGGTTTCGCCATTTGGCCAGGCTGATCTTGAATC	2940
Db	2881	AATTTTGTATTTTAAATATAGAGACGGGGTTTCGCCATTTGGCCAGGCTGATCTTGAATC	2940
QY	2941	CTGACGCTCAAGTAGCTGCTGCTGCTGCTTGGTCTTCCCAATACAGGCATGAAACCATGACCCA	3000
Db	2941	CTGACGCTCAAGTAGCTGCTGCTGCTGCTTGGTCTTCCCAATACAGGCATGAAACCATGACCCA	3000
QY	3001	CCTACTTAGATATTTTCAATGCTATAGACATTTAGAGAGATTTTTCATTTTCCATGACAT	3060
Db	3001	CCTACTTAGATATTTTCAATGCTATAGACATTTAGAGAGATTTTTCATTTTCCATGACAT	3060
QY	3061	TTTTCCTCTCTGCAAAATGGCTTAGCTTGTGTTTTCCTTTTGGGGCAAGACAGAT	3120
Db	3061	TTTTCCTCTCTGCAAAATGGCTTAGCTTGTGTTTTCCTTTTGGGGCAAGACAGAT	3120
QY	3121	CATTAATAATTCGTGACATTTTTCCTTATCAAGAGATATATCAGTGTGTTCTCATGGA	3180
Db	3121	CATTAATAATTCGTGACATTTTTCCTTATCAAGAGATATATCAGTGTGTTCTCATGGA	3180
QY	3181	ACTGCGCTGATTCATTTATGTTTTTTCGATCCATCCTGTCGCTGATCCCTGATCCCTGAC	3240
Db	3181	ACTGCGCTGATTCATTTATGTTTTTTCGATCCATCCTGTCGCTGATCCCTGATCCCTGAC	3240
QY	3241	TCCTTTGGTATTTTCACTGAATTTCAACAATTTTGTGAGAGAGAGAGAGAGAGAGAGAG	3300
Db	3241	TCCTTTGGTATTTTCACTGAATTTCAACAATTTTGTGAGAGAGAGAGAGAGAGAGAGAG	3300
QY	3301	GAAG	3360
Db	3301	GAAG	3360
QY	3361	GAAG	3420
Db	3361	GAAG	3420
QY	3421	TCCTTTGGTATTTTCACTGAATTTCAACAATTTTGTGAGAGAGAGAGAGAGAGAGAGAG	3480
Db	3421	TCCTTTGGTATTTTCACTGAATTTCAACAATTTTGTGAGAGAGAGAGAGAGAGAGAGAG	3480
QY	3481	GAAG	3540
Db	3481	GAAG	3540
QY	3541	GAAG	3600
Db	3541	GAAG	3600
QY	3601	GAAG	3660
Db	3601	GAAG	3660
QY	3661	GAAG	3720
Db	3661	GAAG	3720
QY	3721	GAAG	3780
Db	3721	GAAG	3780
QY	3781	GAAG	3840
Db	3781	GAAG	3840
QY	3841	GAAG	3900
Db	3841	GAAG	3900
QY	3901	GAAG	3960
Db	3901	GAAG	3960
QY	3961	GAAG	4020
Db	3961	GAAG	4020
QY	4021	GAAG	4080
Db	4021	GAAG	4080
QY	4081	GAAG	4140
Db	4081	GAAG	4140
QY	4141	GAAG	4200
Db	4141	GAAG	4200
QY	4201	GAAG	4260
Db	4201	GAAG	4260
QY	4261	GAAG	4320
Db	4261	GAAG	4320
QY	4321	GAAG	4380
Db	4321	GAAG	4380
QY	4381	GAAG	4440
Db	4381	GAAG	4440
QY	4441	GAAG	4500
Db	4441	GAAG	4500
QY	4501	GAAG	4560
Db	4501	GAAG	4560
QY	4561	GAAG	4620
Db	4561	GAAG	4620
QY	4621	GAAG	4680
Db	4621	GAAG	4680
QY	4681	GAAGAGAGAGAG	

Query Match	Best Local Similarity	Score	31.4, 8	DB 24	Length 3654
Matches 3119	Conservative	0	Mismatches	7	Indels
					Gaps
					0
XX	The invention relates to a novel method for diagnosing and detecting the				
CC	progression of liver cancer, hepatocellular carcinoma or metastatic liver				
CC	tumour in a patient, and differentiating metastatic liver cancer from				
CC	hepatocellular carcinoma in a patient, involving detecting the level of				
CC	expression of two or more genes represented in ABR93503-ABR97455 in a				
CC	tissue sample. The method of the invention has hepatocytotropic, and				
CC	cytostatic activity. The method is useful for diagnosing and detecting				
CC	the progression of liver cancer, hepatocellular carcinoma and metastatic				
CC	liver carcinoma in a patient. The method is useful for identifying				
CC	expression profiles which serve as useful diagnostic markers as well as				
CC	markers that can be used to monitor disease states, disease progression,				
CC	drug toxicity, drug efficacy and drug metabolism.				
CC	Note: The sequence data for this patent did not form part of the printed				
CC	specification, but was obtained in electronic format directly from WIPO				
CC	at Etp.wipo.int/pub/published_pat_sequences.				
XX					
XX					
SQ	Sequence 3654 BP, 1052 A; 804 C; 777 G; 1021 T; 0 other;				
Query Match	99.1%; Score 3314.8; DB 24; Length 3654;				
Best Local Similarity	99.8%; Pred. No. 0;				
Matches 3119	Conservative	0	Mismatches	7	Indels
					Gaps
					0
QY	8 GTCCTGACCACTGAATGGAAGAAAAGACTTTTAAACCACTTTGTGACTTACAGAAAG	67			
DB	13 GTCTCGACCACTGAATGGAAGAAAAGACTTTTAAACCACTTTGTGACTTACAGAAAG	72			
QY	68 GAATTTGAAATAAGAAACCTATGATTACTTCAGGCCCATCTTCACTCCCTGTGCTTCTTA	127			
DB	73 GAATTTGAAATAAGAAACCTATGATTACTTCAGGCCCATCTTCACTCCCTGTGCTTCTTA	132			
QY	128 TGCTTTATTTGGCAACTGGATATGATGCGCAAGGGGAAAGTTAGTGAACCCCTGAAACCA	187			
DB	133 TGCTTTATTTGGCAACTGGATATGATGCGCAAGGGGAAAGTTAGTGAACCCCTGAAACCA	192			
QY	188 TGACATTTTCTATTATATGAGAGCCAGATCGAATCAATTAATTCAGTTTAAGGCCA	247			
DB	193 TGACATTTTCTATTATATGAGAGCCAGATCGAATCAATTAATTCAGTTTAAGGCCA	252			
QY	248 ATCTCTCTGCTGTGACTTTGAACCTAAGGGGAGACAGACAACTATTTGTGATAGAAC	307			
DB	253 ATCTCTCTGCTGTGACTTTGAACCTAAGGGGAGAGACAGACAACTATTTGTGATAGAAC	312			
QY	308 GGGAGGGACCTCTGTATTAACAACAGAGCTTGGACAGGGAACAAGATCTTCTACATC	367			
DB	313 GGGAGGGACCTCTGTATTAACAACAGAGCTTGGACAGGGAACAAGATCTTCTACATC	372			
QY	368 TCCAGGTTGACGCCCTGAGCCCTATATGATATATGAGAGGTCAGATCCCTATCAACA	427			
DB	373 TCCAGGTTGACGCCCTGAGCCCTATATGATATATGAGAGGTCAGATCCCTATCAACA	432			
QY	428 TAGAAGTAGAGGACATCAACGACATGACCCACGTTTCTCAGTCAAAAGTACGAAGCT	487			
DB	433 TAGAAGTAGAGGACATCAACGACATGACCCACGTTTCTCAGTCAAAAGTACGAAGCT	492			
QY	488 CAGTAGGACGAACCTCTGCGCCAGAGAAAGCCCTCTTGTATGTCAAGTCCACAGCCCTG	547			
DB	493 CAGTAGGACGAACCTCTGCGCCAGAGAAAGCCCTCTTGTATGTCAAGTCCACAGCCCTG	552			
QY	548 ATGATCCGGGCACTCCCAATGGCAGCTTTATTAACAGATGTGATCCAGCTTCCATGA	607			
DB	553 ATGATCCGGGCACTCCCAATGGCAGCTTTATTAACAGATGTGATCCAGCTTCCATGA	612			
QY	608 TCAACAATGTCACTACTTTCAGATCAACAACAAGGGAGCCATCTCTTACCCGAG	667			
DB	613 TCAACAATGTCACTACTTTCAGATCAACAACAAGGGAGCCATCTCTTACCCGAG	672			
QY	668 AGGAGATTCAGGAATTGAATCTCTGTAAAGATCTCTTCTATTAATCTGGTATCTCAGTGA	727			
DB	673 AGGAGATTCAGGAATTGAATCTCTGTAAAGATCTCTTCTATTAATCTGGTATCTCAGTGA	732			
QY	728 AGGAGATGGAGGCCAGATGGAATTCCTCAGTATACCAACATCTGTGATATCAATAG	787			

Db 733 AGGACATGGAGGCCAAGATGAGAAATTCCTTCACTGATACCAATCTGTGGATATATCAATAG 792
Qy 788 TGAACAGAAATATTTGGAAAGCACAAACCTGTGAGATGTGGAAAACTCAACTGATC 847
Db 793 TGACAGAGAAATATTTGGAAAGCACAAACCTGTGAGATGTGGAAAACTCAACTGATC 852
Qy 848 CTCACCCCATCAAAATCACTCAGGTGGGTGGAAATGATCCCGGTGCAAAATATTCCTTAG 907
Db 853 CTCACCCCATCAAAATCACTCAGGTGGGTGGAAATGATCCCGGTGCAAAATATTCCTTAG 912
Qy 908 TTGACAAAGAGAGCTGCCAAGANTCCCATTTTCAATTTGACCGAGAAAGATATTTACG 967
Db 913 TTGACAAAGAGAGCTGCCAAGANTCCCATTTTCAATTTGACCGAGAAAGATATTTACG 972
Qy 968 TGACTGAGCCCTTGGACCGAGAGAAAGAGATGATGTTTTTTATGCACTTGCAAGG 1027
Db 973 TGACTGAGCCCTTGGACCGAGAGAAAGAGATGATGTTTTTTATGCACTTGCAAGG 1032
Qy 1028 ATGAGTACGGAAAAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTAAAGATATTA 1087
Db 1033 ATGAGTACGGAAAAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTAAAGATATTA 1092
Qy 1088 ATGATATATCCACTCACTACATGTCCTGTCACACAGTAAACCGTATTTGAGTCCAGAGATGAA 1147
Db 1093 ATGATATATCCACTCACTACATGTCCTGTCACACAGTAAACCGTATTTGAGTCCAGAGATGAA 1152
Qy 1148 GACTGGGTAAACAGTATCGGAGACCCCTTACTGACATGACAGGAGTGAAGAAATATCTGCCA 1207
Db 1153 GACTGGGTAAACAGTATCGGAGACCCCTTACTGACATGACAGGAGTGAAGAAATATCTGCCA 1212
Qy 1208 ACAGTTTTCTTAAACCTACAGAGATGTGGAGCAAACTCCCAAACTTCCCATGAGATGACTCT 1267
Db 1213 ACAGTTTTCTTAAACCTACAGAGATGTGGAGCAAACTCCCAAACTTCCCATGAGATGACTCT 1272
Qy 1268 TCCATATCCAAACCTATGCTGGAATGTATACGTTAGTTAAACAGTCCCTTGAAGAGCAG 1327
Db 1273 TCCATATCCAAACCTATGCTGGAATGTATACGTTAGTTAAACAGTCCCTTGAAGAGCAG 1332
Qy 1328 ATACTCCTCAGTACCACTTAACGATAGAGGTGTCTGACAAAGATTTCAAGACCTTTGTT 1387
Db 1333 ATACTCCTCAGTACCACTTAACGATAGAGGTGTCTGACAAAGATTTCAAGACCTTTGTT 1392
Qy 1388 TTGTGCAATCAACGTTATTTGATATCAATGATCAGATCCCATCTTTGAAAAATCAAGTT 1447
Db 1393 TTGTGCAATCAACGTTATTTGATATCAATGATCAGATCCCATCTTTGAAAAATCAAGTT 1452
Qy 1448 ATGGAACCTGATCTTTGCTGGAAGACAAACCTTGGGTCCACATCTTTAACATCCACG 1507
Db 1453 ATGGAACCTGATCTTTGCTGGAAGACAAACCTTGGGTCCACATCTTTAACATCCACG 1512
Qy 1508 CCACTGATGCTGATGAGCCATTACTGGGAGTTCTAAAAATTCGTATNCATATCATTAAG 1567
Db 1513 CCACTGATGCTGATGAGCCATTACTGGGAGTTCTAAAAATTCGTATNCATATCATTAAG 1572
Qy 1568 GAGACAGTGAAGGAGCCCTGGGGGTTGACACAGATCCCATCAACACCGGATATGTCA 1627
Db 1573 GAGACAGTGAAGGAGCCCTGGGGGTTGACACAGATCCCATCAACACCGGATATGTCA 1632
Qy 1628 TAATTTAAAAAGCTCTTTGATTTTGAACACAGACGTGTTTCAACATGTGTCTAAAGCAG 1687
Db 1633 TAATTTAAAAAGCTCTTTGATTTTGAACACAGACGTGTTTCAACATGTGTCTAAAGCAG 1692
Qy 1688 AAAATCTGAGCTCTAGTGTGTTGTTGATGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1747
Db 1693 AAAATCTGAGCTCTAGTGTGTTGTTGATGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1752
Qy 1748 CGCTTATTTGACAGATGTGAATGAAGCACTCAATTTTCCCAACAGTATTTCAAGCA 1807
Db 1753 CGCTTATTTGACAGATGTGAATGAAGCACTCAATTTTCCCAACAGTATTTCAAGCA 1812
Qy 1808 AAGTCAGTGAAGTATGATCTATATGGAATTAAGTGGGCAATGTGATCGCAAGATCCAG 1867
Db 1813 AAGTCAGTGAAGTATGATCTATATGGAATTAAGTGGGCAATGTGATCGCAAGATCCAG 1872

Qy 1868 AAGTCTGACATAAGCTATATCTAGAGGGGAGACACAAGAGGTGGCTTAAATTTGACC 1927
Db 1873 AAGTCTGACATAAGCTATATCTAGAGGGGAGACACAAGAGGTGGCTTAAATTTGACC 1932
Qy 1928 ACGTACCTGTGAGATCTTTAGTGTGCTCCATTTGACAGAGAAAGCCGAATTCATATC 1987
Db 1933 ACGTACCTGTGAGATCTTTAGTGTGCTCCATTTGACAGAGAAAGCCGAATTCATATC 1992
Qy 1988 GGGTACAAAGTGTGGCCACAGAGTAAAGGGGGCTTCTTAAGCTCTGTGTCAAGATTCC 2047
Db 1993 GGGTACAAAGTGTGGCCACAGAGTAAAGGGGGCTTCTTAAGCTCTGTGTCAAGATTCC 2052
Qy 2048 ACCGTATCTTATGATATGATATGAACAACCTCCAGAGCTAGCCAGGACTACACGGGCT 2107
Db 2053 ACCGTATCTTATGATATGATATGAACAACCTCCAGAGCTAGCCAGGACTACACGGGCT 2112
Qy 2108 TGTCTTCTGCAATCCCTCAGTGACCTGGAAGTCAATTTTGAAGGCTACTGATATG 2167
Db 2113 TGTCTTCTGCAATCCCTCAGTGACCTGGAAGTCAATTTTGAAGGCTACTGATATG 2172
Qy 2168 ATACGACTTATTTGGGGGTCCCATTTTACATTTTCCCTGGCAGTGAAGCTTACAA 2227
Db 2173 ATACGACTTATTTGGGGGTCCCATTTTACATTTTCCCTGGCAGTGAAGCTTACAA 2232
Qy 2228 ACGACTGGGAAGTTTCCAAATCAATGTAATGTAATCCGCACTGTCTACACGGCACAG 2287
Db 2233 ACGACTGGGAAGTTTCCAAATCAATGTAATGTAATCCGCACTGTCTACACGGCACAG 2292
Qy 2288 ACTTTGAGAGAGAGGCGTATGTCTGCTTATGATCCGATCATATGAGGGGTCCGCCACCT 2347
Db 2293 ACTTTGAGAGAGAGGCGTATGTCTGCTTATGATCCGATCATATGAGGGGTCCGCCACCT 2352
Qy 2348 TGGAGGCAATGTTTCTTTTACAGTTTACATTTCTGCAAGTGTGTGAGAGAGTGTTC 2407
Db 2353 TGGAGGCAATGTTTCTTTTACAGTTTACATTTCTGCAAGTGTGTGAGAGAGTGTTC 2412
Qy 2408 GGCAGCAGAGTCAACGACCTGGGATACCCACTGTGGGCAATGGCAGTGTATATACGCTGA 2467
Db 2413 GGCAGCAGAGTCAACGACCTGGGATACCCACTGTGGGCAATGGCAGTGTATATACGCTGA 2472
Qy 2468 CCAACCTTCTGATGTTGATATATTTTATGAGATGTGTATTTTACCGCATTAAGAAAGATA 2527
Db 2473 CCAACCTTCTGATGTTGATATATTTTATGAGATGTGTATTTTACCGCATTAAGAAAGATA 2532
Qy 2528 AAGGCAAGATATGTTGAAAGTGTCAAAGCATGTGAAGTCAAACCTCTGAAGAGCTGAA 2587
Db 2533 AAGGCAAGATATGTTGAAAGTGTCAAAGCATGTGAAGTCAAACCTCTGAAGAGCTGAA 2592
Qy 2588 TTTGAAAAGGAATGTTTGAATTTATATGCAAGTGTCTATTTGAGCAACACATCTCATC 2647
Db 2593 TTTGAAAAGGAATGTTTGAATTTATATGCAAGTGTCTATTTGAGCAACACATCTCATC 2652
Qy 2648 CTATTAATTTTCAATCAAGTGCATTATATTTTAAACAGATATTCCTCTTGTCTT 2707
Db 2653 CTATTAATTTTCAATCAAGTGCATTATATTTTAAACAGATATTCCTCTTGTCTT 2712
Qy 2708 TAATATTTGCTAAATATTTCTTTTGAAGTGAAGTCTTGTCTGTGCCAGGCTGAG 2767
Db 2713 TAATATTTGCTAAATATTTCTTTTGAAGTGAAGTCTTGTCTGTGCCAGGCTGAG 2772
Qy 2768 TACAGTGTGATATCCAGCTCACTGCAACCTCCGCTCTCTGGGTTCAATATCTCTCT 2827
Db 2773 TACAGTGTGATATCCAGCTCACTGCAACCTCCGCTCTCTGGGTTCAATATCTCTCT 2832
Qy 2828 GCTCAGCTTCTTAAGTGTGGGTTTACAGGCAACCAACCATGAGCCAGCTAAATTTT 2887
Db 2833 GCTCAGCTTCTTAAGTGTGGGTTTACAGGCAACCAACCATGAGCCAGCTAAATTTT 2892
Qy 2888 GTATTTTAAATAGACAGGGGTTTGGCAATTTGGCAGGCTGTGCTTGAATCTCTGACGT 2947
Db 2893 GTATTTTAAATAGACAGGGGTTTGGCAATTTGGCAGGCTGTGCTTGAATCTCTGACGT 2952

DG 193 TGACATTTTCTATTGTAAGGCCAAGAACCGAGTCAATTAATATTCAGTTTAAGGCCA 252
QY 248 ATCTCTGCTGTGACTTTTGAACTAATCTGGGAGACAGACACAATATTTGTATAGAAC 307
DB 253 ATCTCTGCTGTGACTTTTGAACTAATCTGGGAGACAGACACAATATTTGTATAGAAC 312
QY 308 GGGAGGAGCTTCTGTATTATCAACAGAGCTTTGACAGAGAAAACAAGTCTACACAAATC 367
DB 313 GGGAGGAGCTTCTGTATTATCAACAGAGCTTTGACAGAGAAAACAAGTCTACACAAATC 372
QY 368 TCCAGGTTGAGCCCTGAGAGCTTAATGAAATTAATGAGAGGCTCCATTCACCA 427
DB 373 TCCAGGTTGAGCCCTGAGAGCTTAATGAAATTAATGAGAGGCTCCATTCACCA 432
QY 428 TAGAAGTGAAGAGCATCAACAGCAATGACCCAGTTTCTCCAGTCAAAAGTACAGAGCT 487
DB 433 TAAAGTGAAGAGCATCAACAGCAATGACCCAGTTTCTCCAGTCAAAAGTACAGAGCT 492
QY 488 CAGTAAAGGAGAGCTCTGAGCCAGAGAAAGCCCTTCTGTATGTCAATGCCACAGAGCTGG 547
DB 493 CAGTAAAGGAGAGCTCTGAGCCAGAGAAAGCCCTTCTGTATGTCAATGCCACAGAGCTGG 552
QY 548 ATGATCCGGGCACTCCCAATGSCAGCTTTATACAGATGTGATGATGATGATGATGATGAT 607
DB 553 ATGATCCGGGCACTCCCAATGSCAGCTTTATACAGATGTGATGATGATGATGATGATGATGAT 612
QY 608 TCAACAATGTCAATGATCTTTCAGATCAACAAACAAAGGAGCCATCTCTTACCCGAG 667
DB 613 TCAACAATGTCAATGATCTTTCAGATCAACAAACAAAGGAGCCATCTCTTACCCGAG 672
QY 668 AGGAGTCTCAAGAAATGAATCTGCTAAGAAATCTTCTCTATATCTGTGTATCTCAAGTGA 727
DB 673 AGGAGTCTCAAGAAATGAATCTGCTAAGAAATCTTCTCTATATCTGTGTATCTCAAGTGA 732
QY 728 AGGAGTCTCAAGAAATGAATCTGCTAAGAAATCTTCTCTATATCTGTGTATCTCAAGTGA 787
DB 733 AGGAGTCTCAAGAAATGAATCTGCTAAGAAATCTTCTCTATATCTGTGTATCTCAAGTGA 792
QY 788 TGAACAGAGATATTTGGAAGACCAAAACCTGTGAGATGTGTGGAAGAACTCAACATGATC 847
DB 793 TGAACAGAGATATTTGGAAGACCAAAACCTGTGAGATGTGTGGAAGAACTCAACATGATC 852
QY 848 CTGACCCCATCAAAATCACTCAAGGTGCGGTGATGATGATGATGATGATGATGATGATGATGAT 907
DB 853 CTGACCCCATCAAAATCACTCAAGGTGCGGTGATGATGATGATGATGATGATGATGATGATGAT 912
QY 908 TTGACAAAGAGAGAGCTGCGCAAGATTTCCATTTTCAATTTACAGAGAGAGATATTTAGC 967
DB 913 TTGACAAAGAGAGAGCTGCGCAAGATTTCCATTTTCAATTTACAGAGAGAGATATTTAGC 972
QY 968 TGACTCAGCCCTTGAACCGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1027
DB 973 TGACTCAGCCCTTGAACCGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1032
QY 1028 ATGAGTACGAG 1087
DB 1033 ATGAGTACGAG 1092
QY 1088 ATGATATATCACTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1147
DB 1093 ATGATATATCACTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1152
QY 1148 GACTGGGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1207
DB 1153 GACTGGGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1212
QY 1208 ACAGTTTTCTTAACTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1267
DB 1213 ACAGTTTTCTTAACTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
QY 1268 TCTTAATCCAACTATGCTGGAATGTTTACAGTTAGTTAAACAGTCTTGAAGAGCAAG 1327
DB 1273 TCTTAATCCAACTATGCTGGAATGTTTACAGTTAGTTAAACAGTCTTGAAGAGCAAG 1332

QY 1328 ATACTCTCAGTACAACTTAACATAGAGGTGTCTGACAAAGATTTCAAGACCCCTTGT 1387
DB 1333 ATACTCTCAGTACAACTTAACATAGAGGTGTCTGACAAAGATTTCAAGACCCCTTGT 1392
QY 1388 TTGTGCAATCAACGTTATTGATATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1447
DB 1393 TTGTGCAATCAACGTTATTGATATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1452
QY 1448 ATGGAACCTGATCTTGTGAGAGACAAACATTTGGGTCCACATCTTAAACATCCAGG 1507
DB 1453 ATGGAACCTGATCTTGTGAGAGACAAACATTTGGGTCCACATCTTAAACATCCAGG 1512
QY 1508 CCACTGATGCTGATGAGAGACATTTACTGAGAGTTCTTAAATTTCTGTATCATATCAATAAG 1567
DB 1513 CCACTGATGCTGATGAGAGACATTTACTGAGAGTTCTTAAATTTCTGTATCATATCAATAAG 1572
QY 1568 GAGACAGTGAAG 1627
DB 1573 GAGACAGTGAAG 1632
QY 1628 TAAATTAAG 1687
DB 1633 TAAATTAAG 1692
QY 1688 AAAATCCTGAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1747
DB 1693 AAAATCCTGAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1752
QY 1748 CGCTTATTTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1807
DB 1753 CGCTTATTTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1812
QY 1808 AAGTCAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1867
DB 1813 AAGTCAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1872
QY 1868 AAGTCTGAGCATTAAGATTTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1927
DB 1873 AAGTCTGAGCATTAAGATTTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1932
QY 1928 ACGTGAAGTGAAGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1987
DB 1933 ACGTGAAGTGAAGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1992
QY 1988 GGGTACAGTGTGTGAG 2047
DB 1993 GGGTACAGTGTGTGAG 2052
QY 2048 ACGTGAAGTGAAGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2107
DB 2053 ACGTGAAGTGAAGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2112
QY 2108 TGTCTTCTGCAATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2167
DB 2113 TGTCTTCTGCAATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2172
QY 2168 ATCAGCATTTATTTGGGGGTCCCATTTTACATTTTCCCTGAGAGAGAGAGAGAGAGAGAGAG 2227
DB 2173 ATCAGCATTTATTTGGGGGTCCCATTTTACATTTTCCCTGAGAGAGAGAGAGAGAGAGAGAG 2232
QY 2228 ACGACGTGGAGAGTTTCAAAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2287
DB 2233 ACGACGTGGAGAGTTTCAAAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2292
QY 2288 ACTTTGAG 2347
DB 2293 AGTTTGAAG 2352
QY 2348 TGAAGAGCATTTGTTTCTTTTACAGATTTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2407
DB 2353 TGAAGAGCATTTGTTTCTTTTACAGATTTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2412

QY 2408 GGCACAGAGTCAACAGACTGGGATACCACTGTGGGAGTGGATGATATCTGCTGA 2467
 |||||
 DB 2413 GGCACAGAGTCAACAGACTGGGATACCACTGTGGGAGTGGATGATATCTGCTGA 2472
 |||||
 QY 2468 CCACCCCTTGGGATGATATATTTAGCAGTGTGTTATCCGCATTAAGAGATA 2527
 |||||
 DB 2473 CCACCCCTTGGGATGATATATTTAGCAGTGTGTTATCCGCATTAAGAGATA 2532
 |||||
 QY 2528 AAGCAAAAGATATGTTGAAGTGTCAAGCATCTGAAGTCAAACTCTGAAGAGCTGA 2587
 |||||
 DB 2533 AAGCAAAAGATATGTTGAAGTGTCAAGCATCTGAAGTCAAACTCTGAAGAGCTGA 2592
 |||||
 QY 2588 TTGAAAAGAAATGTTGAATTTATATAGCAAGTCAATTTCCAGCAAAACCATCTCATC 2647
 |||||
 DB 2593 TTGAAAAGAAATGTTGAATTTATATAGCAAGTCAATTTCCAGCAAAACCATCTCATC 2652
 |||||
 QY 2648 CTATTAATTTCTCACTCAAGCTGATTAATTTTAAAGATATTCCTCTTGTCTT 2707
 |||||
 DB 2653 CTATTAATTTCTCACTCAAGCTGATTAATTTTAAAGATATTCCTCTTGTCTT 2712
 |||||
 QY 2708 TAAATATTTGCTAAATATTTCTTTTGAAGTGAAGTGTCTGTGCTGAGCTGAG 2767
 |||||
 DB 2713 TAAATATTTGCTAAATATTTCTTTTGAAGTGAAGTGTCTGTGCTGAGCTGAG 2772
 |||||
 QY 2768 TACAGTGTGTGATCCAGCTCACTGCAACTCCGCTCTGAGGTTCAATGATTCCT 2827
 |||||
 DB 2773 TACAGTGTGTGATCCAGCTCACTGCAACTCCGCTCTGAGGTTCAATGATTCCT 2832
 |||||
 QY 2828 GCCCTACGCTCTCAAGTGTGAGTGGTTTACAGGACCAACCAACATGCGCACTAATTTT 2887
 |||||
 DB 2833 GCCCTACGCTCTCAAGTGTGAGTGGTTTACAGGACCAACCAACATGCGCACTAATTTT 2892
 |||||
 QY 2888 GATATTTTAAATAGAGACGGGGTTTCCCATTTGGCCAGGCTGTCTTGAATCTCTGACGT 2947
 |||||
 DB 2893 GATATTTTAAATAGAGACGGGGTTTCCCATTTGGCCAGGCTGTCTTGAATCTCTGACGT 2952
 |||||
 QY 2948 CAAATGATCTGCTGCTGCTTGTGTCTCCCAATACAGGCAATGAAACACTGACCACTTCTT 3007
 |||||
 DB 2953 CAAATGATCTGCTGCTGCTTGTGTCTCCCAATACAGGCAATGAAACACTGACCACTTCTT 3012
 |||||
 QY 3008 AGATATTTCAATGTCTATAGACATTTAGAGATTTTTCATTTTCCATGACATTTTCT 3067
 |||||
 DB 3013 AGATATTTCAATGTCTATAGACATTTAGAGATTTTTCATTTTCCATGACATTTTCT 3072
 |||||
 QY 3068 CTCTGCAAAATGCTTAACTGTGTGTTTCCCTTTGGGCAAGACAGACTATTAA 3127
 |||||
 DB 3073 CTCTGCAAAATGCTTAACTGTGTGTTTCCCTTTGGGCAAGACAGACTATTAA 3132
 |||||
 QY 3128 TATCTGTACATTTTCTTATATCAAGAGATATATCAAGTGTGTCTCATGAACTGCT 3187
 |||||
 DB 3133 TATCTGTACATTTTCTTATATCAAGAGATATATCAAGTGTGTCTCATGAACTGCT 3192
 |||||
 QY 3188 GGATTCATTTATGTTTCTTATGATTCATCTGTGCTCCCTTCACTCTTGG 3247
 |||||
 DB 3193 GGATTCATTTATGTTTCTTATGATTCATCTGTGCTCCCTTCACTCTTGG 3252
 |||||
 QY 3248 GTATTTCACTGAATTTCAAACTTTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3307
 |||||
 DB 3253 GTATTTCACTGAATTTCAAACTTTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3312
 |||||
 QY 3308 AAAATNAATNAAG 3333
 |||||
 DB 3313 AAAATNAATNAAG 3338
 |||||

DE Oesophagus cancer related gene sequence SEQ ID NO:6004.
 XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;
 gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 XX
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 20-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.
 XX
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set
 XX
 XX Claim 1; SEQ ID 6004; 44pp; English.
 PS
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical

CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in AB161664
CC to AB170110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophagal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.

XX Sequence 3654 BP; 1052 A; 804 C; 777 G; 1021 T; 0 other;

Query Match 99.1%; Score 3314.8; DB 24; Length 3654;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3319; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 8 GTCTCGACCACTGAATGGAAGAAAGAGCTTTTACCACTTTTGACTTACAGAAAG 67
DB 13 GTCTCGACCACTGAATGGAAGAAAGAGCTTTTAAACCACTTTTGACTTACAGAAAG 72
QY 68 GAATTTGAATTAAGAAAGAACTATGATCTTCAGAGCCCATCTCCCTGTCCTTTA 127
DB 73 GAATTTGAATTAAGAAAGAACTATGATCTTCAGAGCCCATCTCCCTGTCCTTTA 132
QY 128 TGCCTTAATTTGGCACTGATATGCGCAAGAGGGGAAGTTAGTGAACCCCTGAAACCA 187
DB 133 TGCCTTAATTTGGCACTGATATGCGCAAGAGGGGAAGTTAGTGAACCCCTGAAACCA 192
QY 188 TGCATTTTCTATTTATGAAGGCCAAGAACCGAGTCAATTTATTTCAAGTTTAAAGCCA 247
DB 193 TGCATTTTCTATTTATGAAGGCCAAGAACCGAGTCAATTTATTTCAAGTTTAAAGCCA 252
QY 248 ATCTCTCTGCTGATCTTTTGAAGTAACTAGTGGGAGACAGACATATTTGTGTAGAAC 307
DB 253 ATCTCTCTGCTGATCTTTTGAAGTAACTAGTGGGAGACAGACATATTTGTGTAGAAC 312
QY 308 GGGAGGAGACTTGTATTTACACAGAGCCCTTGGACAGGGAACAAGATCTTACACATC 367
DB 313 GGGAGGAGACTTGTATTTACACAGAGCCCTTGGACAGGGAACAAGATCTTACACATC 372
QY 368 TCCAGGTTGAGCCCTTGGAGGCTTATATGATGAGGGTCCAGTCCCTATCCCA 427
DB 373 TCCAGGTTGAGCCCTTGGAGGCTTATATGATGAGGGTCCAGTCCCTATCCCA 432
QY 428 TAGAAGTGAAGGACATCAACGACATGACCCAGTTCCTCCAGTCAAGTAAAGAGGCT 487
DB 433 TAGAAGTGAAGGACATCAACGACATGACCCAGTTCCTCCAGTCAAGTAAAGAGGCT 492
QY 488 CAGTAAAGGCAAGCTCTCGCCACAGAAAGCCCTTGTATGTCAATGACACAGACCTGG 547
DB 493 CAGTAAAGGCAAGCTCTCGCCACAGAAAGCCCTTGTATGTCAATGACACAGACCTGG 552
QY 548 ATATATCCGGCACTCCCAATGCGCAGCTTTATTAACAGATTTGATCCAGCTTCCATGA 607
DB 553 ATATATCCGGCACTCCCAATGCGCAGCTTTATTAACAGATTTGATCCAGCTTCCATGA 612
QY 608 TCAACATGTCATGTACTTTCAATCAACAAACAAAGGAGCATCTCTTAACCCGAG 667
DB 613 TCAACATGTCATGTACTTTCAATCAACAAACAAAGGAGCATCTCTTAACCCGAG 672
QY 668 AGGAGATCTCAGAAATGAATCTCTGCTAAGATCTTCTTAATCTGTATCTCAAGTA 727
DB 673 AGGAGATCTCAGAAATGAATCTCTGCTAAGATCTTCTTAATCTGTATCTCAAGTA 732
QY 728 AGGACATGGAGGCGCAGAGTGAATTTCTTCAAGTATCCACATCTGTGGATATCAATG 787
DB 733 AGGACATGGAGGCGCAGAGTGAATTTCTTCAAGTATCCACATCTGTGGATATCAATG 792

QY 788 TGACAGAGATATTTGGAAAGACACAAAACCTGTGAGATGTGAAAACTCACTGATC 847
DB 793 TGACAGAGATATTTGGAAAGACACAAAACCTGTGAGATGTGAAAACTCACTGATC 852
QY 848 CTACACCCCATCAAAATCTACTCAGTGTGGTGAATGATCCGGTGCAATATTCCTTAG 907
DB 853 CTACACCCCATCAAAATCTACTCAGTGTGGTGAATGATCCGGTGCAATATTCCTTAG 912
QY 908 TTGACAAAGAGAGCTGCCAAGATTCCTCAATTTTCAATTTGACAGAAAGAGATATTTACG 967
DB 913 TTGACAAAGAGAGCTGCCAAGATTCCTCAATTTTCAATTTGACAGAAAGAGATATTTACG 972
QY 968 TGACTCAGCCCTTGACACCGAGAAAGAGATCATATGTTTTTTATGCAAGTTGCAAGG 1027
DB 973 TGACTCAGCCCTTGACACCGAGAAAGAGATCATATGTTTTTTATGCAAGTTGCAAGG 1032
QY 1028 ATGAGTACGGAAAAACCACTTCATATCCGCTGGAAAATTCATGTAAAGTTAAAGATATTA 1087
DB 1033 ATGAGTACGGAAAAACCACTTCATATCCGCTGGAAAATTCATGTAAAGTTAAAGATATTA 1092
QY 1088 ATGATATCCACTATCATATGTCGTACACAGTAAACCGTATTTGAGTCCAGAGAAATGAC 1147
DB 1093 ATGATATCCACTATCATATGTCGTACACAGTAAACCGTATTTGAGTCCAGAGAAATGAC 1152
QY 1148 GACTGGGTAAACAGTATCGGGACCCCTTACTGACATGACAGGGATGAAGAAAATTAAGTCCA 1207
DB 1153 GACTGGGTAAACAGTATCGGGACCCCTTACTGACATGACAGGGATGAAGAAAATTAAGTCCA 1212
QY 1208 ACAGTTTTCTTAACTACAGAGATTTGTGAGCAAACTCCCAACTTCCATGATGACTCT 1267
DB 1213 ACAGTTTTCTTAACTACAGAGATTTGTGAGCAAACTCCCAACTTCCATGATGACTCT 1272
QY 1268 TCCATATCCAAACCTATGCTGGAATGTACAGTTACTTAAACAGTCCCTGGAAGCAAG 1327
DB 1273 TCCATATCCAAACCTATGCTGGAATGTACAGTTACTTAAACAGTCCCTGGAAGCAAG 1332
QY 1328 ATACTCCTCAGTCAACTTAAGATAGAGTGTCTGACAAAGATTTCAAGACCCCTTGT 1387
DB 1333 ATACTCCTCAGTCAACTTAAGATAGAGTGTCTGACAAAGATTTCAAGACCCCTTGT 1392
QY 1388 TTGTGCAATCAACGTTATTTGATATCAATGATCAGATCCCATCTTTGAAAAATCAGATT 1447
DB 1393 TTGTGCAATCAACGTTATTTGATATCAATGATCAGATCCCATCTTTGAAAAATCAGATT 1452
QY 1448 ATGAAAACCTGACTCTTGTGTAAGACACAAACCTTGGTCCACATCTTAACTCCAGG 1507
DB 1453 ATGAAAACCTGACTCTTGTGTAAGACACAAACCTTGGTCCACATCTTAACTCCAGG 1512
QY 1508 CCACTGATGCTGATGAGCATTTTACGAGAGTTCTTAAATTTGATCATCATATTAAGG 1567
DB 1513 CCACTGATGCTGATGAGCATTTTACGAGAGTTCTTAAATTTGATCATCATATTAAGG 1572
QY 1568 GAGACAGTGAAGGACGCTGTGGGGGTGACACAGATCCCATCAACACCGGATATGTCA 1627
DB 1573 GAGACAGTGAAGGACGCTGTGGGGGTGACACAGATCCCATCAACACCGGATATGTCA 1632
QY 1628 TTAATTAAGGCTCTTGAATTTTGAACAGACGCTGTTTCCAAATTTGTTTCAAGCAG 1687
DB 1633 TTAATTAAGGCTCTTGAATTTTGAACAGACGCTGTTTCCAAATTTGTTTCAAGCAG 1692
QY 1688 AAAATCCTGAGCCTCAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1747
DB 1693 AAAATCCTGAGCCTCAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1752
QY 1748 CGCTTATTTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1807
DB 1753 CGCTTATTTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1812
QY 1808 AAGTCAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1867
DB 1813 AAGTCAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1872

QY 1868 AAGCTGTGACATTAAGCTATTCACTGAGGGGAGACAGAGAGTGTGGCTTAAATTAAC 1927
 DB 1873 AAGGTCTGAGACATTAAGCTATTCACTGAGGGGAGACAGAGAGTGTGGCTTAAATTAAC 1932
 QY 1928 ACGTGAATGTGAGATCTTTAGTGTGCTCCATTGGAAGAAGCCGGAAGTCCATATC 1987
 DB 1933 ACGTGAATGTGAGATCTTTAGTGTGCTCCATTGGAAGAAGCCGGAAGTCCATATC 1992
 QY 1988 GGGTACAAAGTGTGAGCCAGAGAGTGTGGGGGTCTTCTTAAGCTCTGTGTCAAGTTCC 2047
 DB 1993 GGGTACAAAGTGTGAGCCAGAGAGTGTGGGGGTCTTCTTAAGCTCTGTGTCAAGTTCC 2052
 QY 2048 ACCGATCCCTTATGATGTGAATGACAACCTCCAGGCTTACCAAGAGTACAGGAGCT 2107
 DB 2053 ACCGATCCCTTATGATGTGAATGACAACCTCCAGGCTTACCAAGAGTACAGGAGCT 2112
 QY 2108 TGTTCCTCTGCGCATCCCTCAGTGCACCTGGAAGTCAATTTTGAAGGCTAATGATG 2167
 DB 2113 TGTTCCTCTGCGCATCCCTCAGTGCACCTGGAAGTCAATTTTGAAGGCTAATGATG 2172
 QY 2168 ATCAGCACTTATTTGCGGGTCCCATTTTACATTTTCCCTGCGAGTGAAGCTTACAA 2227
 DB 2173 ATCAGCACTTATTTGCGGGTCCCATTTTACATTTTCCCTGCGAGTGAAGCTTACAA 2232
 QY 2228 ACGACTGGAAAGTTCCAAAATCAATGATCTCATGCCGACTCTACAGGACAG 2287
 DB 2233 ACGACTGGAAAGTTCCAAAATCAATGATCTCATGCCGACTCTACAGGACAG 2292
 QY 2288 ACTTGAAGAGAGGCGGTATGCTGCTTGAATCCGATCAATGATGGGGTCCGACCT 2347
 DB 2293 AGTTGAAGAGAGGAGTATGCTGCTTGAATCCGATCAATGATGGGGTCCGACCT 2352
 QY 2348 TGAAGAGCATTTTCTTTACAGTTACATCTGCAAGTGTGGAAGGAAGTTGTTCC 2407
 DB 2353 TGAAGAGCATTTTCTTTACAGTTACATCTGCAAGTGTGGAAGGAAGTTGTTCC 2412
 QY 2408 GGCCAGAGGTCACACAGCTGGATACCCATCTGGGAGTGGAGTGTATCTGCTGA 2467
 DB 2413 GGCCAGAGGTCACACAGCTGGATACCCATCTGGGAGTGGAGTGTATCTGCTGA 2472
 QY 2468 CCACCTTCTGTGATTTGTATTAATTTTGAAGTGTGTTTATCCGATAAAGAGATA 2527
 DB 2473 CCACCTTCTGTGATTTGTATTAATTTTGAAGTGTGTTTATCCGATAAAGAGATA 2532
 QY 2528 AAGCAAAAGATTAATTTGGAAGTGTCTCAAGATCTGAAGTAAACCTCTGAAGCTGAA 2587
 DB 2533 AAGCAAAAGATTAATTTGGAAGTGTCTCAAGATCTGAAGTAAACCTCTGAAGCTGAA 2592
 QY 2588 TTTGAAAAGGAATTTGGAATTTATATAGCAAGTGTCTTCAAGCAACACATCTCATC 2647
 DB 2593 TTTGAAAAGGAATTTGGAATTTATATAGCAAGTGTCTTCAAGCAACACATCTCATC 2652
 QY 2648 CTATTACTTTTCACTAAGCTGATTAATTTTAAACAGATATTCCTCTGTGCTT 2707
 DB 2653 CTATTACTTTTCACTAAGCTGATTAATTTTAAACAGATATTCCTCTGTGCTT 2712
 QY 2708 TAAATTTTGTAAATTTTCTTTTGAAGTGTGATCTGTGTGCGCCAGGCTGAG 2767
 DB 2713 TAAATTTTGTAAATTTTCTTTTGAAGTGTGATCTGTGTGCGCCAGGCTGAG 2772
 QY 2768 TACAGTGTGTGATCCAGCTCACTGCACTCCGCTCTGAGGTTCAATGATTTCTT 2827
 DB 2773 TACAGTGTGTGATCCAGCTCACTGCACTCCGCTCTGAGGTTCAATGATTTCTT 2832
 QY 2828 GCCTCAGCTTCTAAGTGTGAGTTTACAGGACCCACACATGCTCCAGCTAATTTT 2887
 DB 2833 GCCTCAGCTTCTAAGTGTGAGTTTACAGGACCCACACATGCTCCAGCTAATTTT 2892
 QY 2888 GTATTTTAAATGAGACGGGGTTCGCAATTTGCGCAGGCTGTGAACTCTGAGCT 2947
 DB 2893 GTATTTTAAATGAGACGGGGTTCGCAATTTGCGCAGGCTGTGAACTCTGAGCT 2952
 QY 2948 CAAGTGAATCTGCTGCTGTGCTCCCAATACAGGATGAACCACTGACCACTACTT 3007

DB 2953 CAAGTGAATCTGCTGCTGTGCTCCCAATACAGGATGAACCACTGACCCACTACTT 3012
 QY 3008 AGATATTTTCAATGTGCTATAGACATTTAGAGATTTTTCATTTTCCATGACATTTTCC 3067
 DB 3013 AGATATTTTCAATGTGCTATAGACATTTAGAGATTTTTCATTTTCCATGACATTTTCC 3072
 QY 3068 CTCTGCAAAATGGCTTACTGATGTTTTCCTTTTGGGGGAGACAGACTCATTTAA 3127
 DB 3073 CTCTGCAAAATGGCTTACTGATGTTTTCCTTTTGGGGGAGACAGACTCATTTAA 3132
 QY 3128 TATTCGTACATTTTCTTTTATCAAGAGATTAATCAAGTGTGTCTATAGACTGCT 3187
 DB 3133 TATTCGTACATTTTCTTTTATCAAGAGATTAATCAAGTGTGTCTATAGACTGCT 3192
 QY 3188 GGATTCATTTATGTTTTCATTCATCCGTGTCCCTTCATCCCTTGAATCTCTTG 3247
 DB 3193 GGATTCATTTATGTTTTCATTCATCCGTGTCCCTTCATCCCTTGAATCTCTTG 3252
 QY 3248 GTATTTCACTGAATTTCAAAATTTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3307
 DB 3253 GTATTTCACTGAATTTCAAAATTTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3312
 QY 3308 AATTAATAAAG 3333
 DB 3313 AATTAATAAAG 3338
 RESULT 9
 ABL37430
 ID ABL37430 standard; cDNA; 523 BP.
 XX
 AC ABL37430;
 DT 08-APR-2002 (first entry)
 XX
 DE Human colon tumour antigen polynucleotide SEQ ID NO:1019.
 XX
 KW Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
 KW colon tumour metastatic antigen; diagnosis; gene; ss.
 OS Homo sapiens.
 PN W0200196388-A2.
 XX
 PD 20-DEC-2001.
 XX
 PF 08-JUN-2001; 2001MO-US18557.
 XX
 PR 09-JUN-2000; 2000US-210899P.
 PR 20-FEB-2001; 2001US-270216P.
 XX
 PA (CORI-) CORIXA CORP.
 PI Jiang Y, Harlocker SL, Secrist H;
 DR WPI; 2002-114514/15.
 XX
 PT Novel isolated colon tumor polynucleotide differentially expressed in
 PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
 PT useful for inhibiting development of cancer in patient -
 XX
 PS Claim 1; SEQ ID 1019; 105bp; English.
 XX
 CC ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (1)
 CC which were isolated from human colon tumour and colon metastatic tumour
 CC cDNA libraries. (1) have cytostatic activity and can be used in vaccine
 CC production. (1) can be used for stimulating and/or expanding T cells
 CC specific for a tumour protein on contact with the T cells. They are also
 CC useful for inhibiting the development of cancer in a patient. (1) can be
 CC used as probes or primers for nucleic acid hybridisation, for preparing
 CC mutant species primers, or primers for use in genetic constructions. (1)
 CC can be used in the diagnosis of a colon tumour.

XX Sequence 523 BP; 133 A; 107 C; 119 G; 162 T; 2 other;
SQ

Query Match 15.4%; Score 514.2; DB 24; Length 523;
Best Local Similarity 99.4%; Pred. No. 4.5e-127;
Matches 516; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2257 ACTCATGCCCGACTGTCTACACGACACAGACTTTGAGGAGAGGGCGTATGTCGTCTG 2316
DB 1 ACTCATGCCCGACTGTCTACACGACACAGACTTTGAGGAGAGGGCGTATGTCGTCTG 60
QY 2317 ATCCGATCAATGATGGGGGTCCGCCACCTTTGGAAGGCATGTTTCTTTACCAAGTTACA 2376
DB 61 ATCCGATCAATGATGGGGGTCCGCCACCTTTGGAAGGCATGTTTCTTTACCAAGTTACA 120
QY 2377 TTCTGAGTGTGTGGAGAGAGTGTTCGGGACAGAGTCCACGACTGGGATGCC 2436
DB 121 TTCTGAGTGTGTGGAGAGAGTGTTCGGGACAGAGTCCACGACTGGGATGCC 180
QY 2437 ACTGTGGGATGGCAGTGTGTATCTGTGACACACCTTCTGTGATTTGGTAAATTTTA 2496
DB 181 ACTGTGGGATGGCAGTGTGTATCTGTGACACACCTTCTGTGATTTGGTAAATTTTA 240
QY 2497 GCACTTGTGTTATCCGATTAAGAGATTAAGGCAAGATTAATGTTGAAGTCTCAA 2556
DB 241 GCACTTGTGTTATCCGATTAAGAGATTAAGGCAAGATTAATGTTGAAGTCTCAA 300
QY 2557 GCATCTGAATCAAACTCTGAGAGAGTGAATTTGAAAAGATTTGAAATTTATATAG 2616
DB 301 GCATCTGAATCAAACTCTGAGAGAGTGAATTTGAAAAGATTTGAAATTTATATAG 360
QY 2617 CAAGTGTATTTGACAAACACATCTCATCTATTTCTTCACTCAAGTCATATA 2676
DB 361 CAAGTGTATTTGACAAACACATCTCATCTATTTCTTCACTCAAGTCATATA 420
QY 2677 ATTTTAAACAGATATTCCTCTGTCTTTAAATTTGCTAAATATTTCTTTTGTAG 2736
DB 421 ATTTTAAACAGATATTCCTCTGTCTTTAAATTTGCTAAATATTTCTTTTGTAG 480
QY 2737 GTGAGTCTGTCTGTGCGCCAGGCGTGAAGTACAGTG 2775
DB 481 GTGAGTCTGTCTGTGCGCCAGGCGTGAAGTACAGTG 519

RESULT 10
AAA16509
ID AAA16509 standard; DNA; 627 BP.
XX
AC AAA16509;
XX
DT 14-JUN-2000 (first entry)
XX
DE Human colon cancer differentially expressed nucleotide sequence #514.
XX
KW Colon cancer; detect; differential expression; human; treatment;
XX detect mutation; non-invasive diagnostic method; ds.
OS Homo sapiens.
XX
XX WO200012702-A2.
XX
XX 09-MAR-2000.
XX
XX 30-AUG-1999; 99WO-US19424.
XX
XX 31-AUG-1998; 98US-0098639.
XX
XX 27-JAN-1999; 99US-0117393.
XX
XX (FARB) BAYER CORP.
XX
PI Endege WO, Steinmann KE, Aetle JH, Burgess CC, Carroll E,
PI Cactino TJ, Dwivedi P, Ford DM, Lewis ME, Molino GA, Monahan JE,
PI Schlegel R;

XX WPI; 2000-256641/22.
XX

PT Novel nucleic acid and proteins for identifying therapeutic agents
XX useful for treating and diagnosing cancer, especially colon cancer -
XX
XX Claim 16; Page 332; 345pp; English.
PS

CC This sequence represents a human nucleotide sequence which is
CC differentially expressed in colon cancer cells compared to the expression
CC levels in normal cells. The nucleotide sequence can be used as a source
CC of primers and probes. The nucleotide sequence is useful for determining
CC the phenotype of a cell by detecting the differential expression of the
CC sequence relative to a normal cell. The probes derived from the sequence
CC can also be used to determine the phenotype of cells in a sample. Probes
CC and antibodies which hybridise to the nucleotide sequence can also be
CC used to determine the phenotype of a cell. The primers are useful for
CC detecting a mutation in a test nucleotide sequence and also for detecting
CC cancer, preferably colon cancer. Antibodies against the protein encoded
CC by the nucleotide sequence can also be used in a method to detect colon
CC cancer. The diagnostic method is non-invasive and accurate for diagnosing
CC colon cancer at an early stage.

XX Sequence 627 BP; 159 A; 134 C; 146 G; 179 T; 9 other;
SQ

Query Match 15.3%; Score 511.4; DB 21; Length 627;
Best Local Similarity 99.0%; Pred. No. 2.8e-126;
Matches 512; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2254 GGTACTCATGCCCGACTGTCTACACGACACAGACTTTGAGAGAGGGCGTATGTCGTC 2313
DB 1 GGTACTCATGCCCGACTGTCTACACGACACAGACTTTGAGAGAGGGCGTATGTCGTC 60
QY 2314 TTGATCCGATCAATGATGGGGGTCCGCCACCTTTGGAAGGCATGTTTCTTTACCAAGTT 2373
DB 61 TTGATCCGATCAATGATGGGGGTCCGCCACCTTTGGAAGGCATGTTTCTTTACCAAGTT 120
QY 2374 ACATTTGAGTGTGTGAGAGAGTGTTCGGGACAGAGTCCACAGACTGGGATA 2433
DB 121 ACATTTGAGTGTGTGAGAGAGTGTTCGGGACAGAGTCCACAGACTGGGATA 180
QY 2434 CCCACTGTGGGACATGGCAGTGTGTATCTGTGACCAACCTTCTGTGATTTGTATATTT 2493
DB 181 CCCACTGTGGGACATGGCAGTGTGTATCTGTGACCAACCTTCTGTGATTTGTATATTT 240
QY 2494 TTGCACTTGTGTTATCCGATTAAGAGATTAAGGCAAGATTAATGTTGAAGTCT 2553
DB 241 TTGCACTTGTGTTATCCGATTAAGAGATTAAGGCAAGATTAATGTTGAAGTCT 300
QY 2554 CAAGCATCTGAAGTCAAACTCTGAGAGAGTGAATTTGAAAAGATTTGAAATTTATA 2613
DB 301 CAAGCATCTGAAGTCAAACTCTGAGAGAGTGAATTTGAAAAGATTTGAAATTTATA 360
QY 2614 TGCAAGTCTATTTGACAAACACATCTCATCTATTTCTTCACTCAAGTCATATA 2673
DB 361 TGCAAGTCTATTTGACAAACACATCTCATCTATTTCTTCACTCAAGTCATATA 420
QY 2674 ATTAATTTTAAACAGATATTCCTCTGTCTTTAAATTTGCTAAATATTTCTTTT 2733
DB 421 ATTAATTTTAAACAGATATTCCTCTGTCTTTAAATTTGCTAAATATTTCTTTT 480
QY 2734 GAGGTGAGTCTGTCTGTGCGCCAGGCGTGAAGTAC 2770
DB 481 GAGGTGAGTCTGTCTGTGCGCCAGGCGTGAAGTAC 517

RESULT 11
ABL38503
ID ABL38503 standard; cDNA; 501 BP.
XX
AC ABL38503;
XX
DT 08-APR-2002 (first entry)

XX DE Human colon tumour antigen polynucleotide SEQ ID NO:2092.
 XX KM Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
 XX KM colon tumour metastatic antigen; diagnosis; gene; ss.
 XX OS Homo sapiens.
 PN WO200196388-A2.
 XX PD 20-DEC-2001.
 XX PF 08-JUN-2001; 2001WO-US18557.
 XX PR 09-JUN-2000; 2000US-210899P.
 XX PR 20-FEB-2001; 2001US-270216P.
 XX PA (CORI-) CORIXA CORP.
 XX PI Jiang Y, Harlocker SL, Secretist H;
 XX DR WPI; 2002-114514/15.
 PT Novel isolated colon tumor polynucleotide differentially expressed in
 PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
 PT useful for inhibiting development of cancer in patient -
 XX PS Claim 1; SEQ ID 2092; 105pp; English.
 CC ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
 CC which were isolated from human colon tumour and colon metastatic tumor
 CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
 CC production. (I) can be used for stimulating and/or expanding T cells
 CC specific for a tumour protein on contact with the T cells. They are also
 CC useful for inhibiting the development of cancer in a patient. (I) can be
 CC used as probes or primers for nucleic acid hybridisation, for preparing
 CC mutant species primers, or primers for use in genetic constructions. (I)
 CC can be used in the diagnosis of a colon tumour.
 XX SQ Sequence 501 BP; 130 A; 99 C; 110 G; 159 T; 3 other;
 Query Match 14.9%; Score 497.2; DB 24; Length 501;
 Best Local Similarity 99.6%; Pred. No. 1.5e-122;
 Matches 497; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2257 ACTCATGCCGACCTGTCTACACGAGCAGACTTTGAGAGAGGCGGTATGCTCTTG 2316
 DB 1 ACTCATGCCGACCTGTCTACACGAGCAGACTTTGAGAGAGGCGGTATGCTCTTG 60
 QY 2317 ATCCGCATCATGATGAGGGGTCGGCCACCTTTGAGAGGCAATTTCTTTACCACTTACA 2376
 DB 61 ATCCGCATCATGATGAGGGGTCGGCCACCTTTGAGAGGCAATTTCTTTACCACTTACA 120
 QY 2377 TTCTGCAGTTGTGTGAGAAAGATTGTTCCGGCCAGAGAGGTCAACAGCTGGGATACC 2436
 DB 121 TTCTGCAGTTGTGTGAGAAAGATTGTTCCGGCCAGAGAGGTCAACAGCTGGGATACC 180
 QY 2437 ACTGTGGGCATGGAGGATGTATACCTGTGACCAACCTTGTGTGATGTATATTTTA 2496
 DB 181 ACTGTGGGCATGGAGGATGTATACCTGTGACCAACCTTGTGTGATGTATATTTTA 240
 QY 2497 GCAAGTTGTGTTATCCGATTAAGAAGATTAAGCAAGATTAAGTGAAGTCTCAA 2556
 DB 241 GCAAGTTGTGTTATCCGATTAAGAAGATTAAGCAAGATTAAGTGAAGTCTCAA 300
 QY 2557 GCATCTGAAGTCAACCTCTGAGAAAGTGAATTTGAAGAATTTATATATAG 2616
 DB 301 GCATCTGAAGTCAACCTCTGAGAAAGTGAATTTGAAGAATTTATATATAG 360
 QY 2617 CAAAGTCTATTTCAGCAACAACATCTCATCTTACTTTTCACTTAAGCTGATATA 2676
 DB 361 CAAAGTCTATTTCAGCAACAACATCTCATCTTACTTTTCACTTAAGCTGATATA 420

QY 2677 ATTTTAAACAGATATCCCTCTGCTCTTAATATTTGCTAAATATTTCTTTTGG 2736
 DB 421 ATTTTAAACAGATATCCCTCTGCTCTTAATATTTGCTAAATATTTCTTTTGG 480
 QY 2737 GTGGAGTCTTGTCTGTGCG 2755
 DB 481 GTGGAGTCTTGTCTGTGCG 499
 RESULT 12
 ABL36982/c
 ID ABL36982 standard; cDNA; 495 BP.
 XX ABL36982;
 AC ABL36982;
 DT 08-APR-2002 (first entry)
 XX DE Human colon tumour antigen polynucleotide SEQ ID NO:571.
 XX KM Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
 XX KM colon tumour metastatic antigen; diagnosis; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200196388-A2.
 XX PD 20-DEC-2001.
 XX PF 08-JUN-2001; 2001WO-US18557.
 XX PR 09-JUN-2000; 2000US-210899P.
 XX PR 20-FEB-2001; 2001US-270216P.
 XX PA (CORI-) CORIXA CORP.
 XX PI Jiang Y, Harlocker SL, Secretist H;
 XX DR WPI; 2002-114514/15.
 PT Novel isolated colon tumor polynucleotide differentially expressed in
 PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
 PT useful for inhibiting development of cancer in patient -
 XX PS Claim 1; SEQ ID 571; 105pp; English.
 CC ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
 CC which were isolated from human colon tumour and colon metastatic tumor
 CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
 CC production. (I) can be used for stimulating and/or expanding T cells
 CC specific for a tumour protein on contact with the T cells. They are also
 CC useful for inhibiting the development of cancer in a patient. (I) can be
 CC used as probes or primers for nucleic acid hybridisation, for preparing
 CC mutant species primers, or primers for use in genetic constructions. (I)
 CC can be used in the diagnosis of a colon tumour.
 XX SQ Sequence 495 BP; 156 A; 114 C; 93 G; 126 T; 6 other;
 Query Match 14.0%; Score 469.4; DB 24; Length 495;
 Best Local Similarity 98.0%; Pred. No. 4.1e-115;
 Matches 483; Conservative 1; Mismatches 7; Indels 2; Gaps 1;
 QY 2278 AGGACACAGACTTTGAGAGAGGCGGTATGCTCTTATCCGCATCATGATGGGGT 2337
 DB 493 AGGACACAGACTTTGAGAGAGGCGGTATGCTCTTATCCGCATCATGATGGGGT 434
 QY 2338 CGGCAACCTTGGAGGCAATGTTCTTTACAGATTATCTGCAAGTTGTGGAAGA 2397
 DB 433 CGGCAACCTTGGAGGCAATGTTCTTTACAGATTATCTGCAAGTTGTGGAAGA 374
 QY 2398 AGTTGTTCCGGCCAGAGGTCAACAGAC--TGGATATCCCACTGTGGGATGGCAATTG 2455
 DB 373 AGTTGTTCCGGCCAGAGGTCAACAGANCTGGGATTAACCACTGTGGGATGGCAATTG 314

QY 2456 GTATATGCTGACCAACCTCTGTGATGGTATTAATTTAGACAGTGTGTTATCCGCA 2515
DB 313 GTATATGCTGACCAACCTCTGTGATGGTATTAATTTAGACAGTGTGTTATCCGCA 254
QY 2516 TTAAGAAGATTAAGGCAAAATATGTTGAAAGTGTCAAGCATCTGAAGTCAAACTC 2575
DB 253 TTAAGAAGATTAAGGCAAAATATGTTGAAAGTGTCAAGCATCTGAAGTCAAACTC 194
QY 2576 TGAGAAGCTGAATTTGAAAAGAAATGTTGAATTTATATAGCAAGTCTATTTGAGCAAC 2635
DB 193 TGAGAAGCTGAATTTGAAAAGAAATGTTGAATTTATATAGCAAGTCTATTTGAGCAAC 134
QY 2636 AACCATCTCATCTATTACTTTTCATCTAACGTGCATATTAATTTTAAACAGATATTC 2695
DB 133 AACCATCTCATCTATTACTTTTCATCTAACGTGCATATTAATTTTAAACAGATATTC 74
QY 2696 CCTTTGCTCTTTTAATATTTGCTAAATTTCTTTTGTAGGTGAGTCTTGTCTGTG 2755
DB 73 CCTTTGCTCTTTTAATATTTGCTAAATATTTCTTTTGTAGGTGAGTCTTGTCTGTG 14
QY 2756 CCCAGGCTGAGT 2768
DB 13 CCCAGGCTGAGT 1
RESULT 13
ABL36509
ID ABL36509 standard; cDNA; 492 BP.
XX ABL36509;
AC
XX
XX
DT 08-APR-2002 (first entry)
XX
DE Human colon tumour antigen polynucleotide SEQ ID NO:98.
XX
XX Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
KM colon tumour metastatic antigen; diagnosis; gene; ss.
XX
OS Homo sapiens.
XX
XX NO200196388-A2.
PN
XX
XX 20-DEC-2001.
PD
XX
XX 08-JUN-2001; 2001WO-US18557.
PF
XX
XX 09-JUN-2000; 2000US-210899P.
PR
XX 20-FEB-2001; 2001US-270216P.
PR
XX (CORI-) CORIXA CORP.
PA
XX
XX
PI Jiang Y, Harlocker SL, Secrist H;
XX
XX WPI; 2002-114514/15.
DR
XX
XX Novel isolated colon tumor polynucleotide differentially expressed in
PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT useful for inhibiting development of cancer in patient -
XX
XX Claim 1; SEQ ID 98; 105bp; English.
PS
XX
XX ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
CC which were isolated from human colon tumour and colon metastatic tumor
CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
CC production. (I) can be used for stimulating and/or expanding T cells
CC specific for a tumour protein on contact with the T cells. They are also
CC useful for inhibiting the development of cancer in a patient. (I) can be
CC used as probes or primers for nucleic acid hybridisation, for preparing
CC mutant species primers, or primers for use in genetic constructions. (I)
CC can be used in the diagnosis of a colon tumour.
XX
XX Sequence 492 BP; 132 A; 94 C; 109 G; 155 T; 2 other;

Query Match 13.8%; Score 461.6; DB 24; Length 492;
Best Local Similarity 99.0%; Pred. No. 5e-113;
Matches 485; Conservative 1; Mismatches 1; Indels 3; Gaps 2;
QY 2257 ACTCATCCCGACTGTCTTACAGGACACAGACTTTGAGAGAGGGCGTATGCTGTTG 2316
DB 1 ACTCATCCCGACTGTCTTACAGGACACAGACTTTGAGAGAGGGCGTATGCTGTTG 60
QY 2217 ATCCGATCAATGATGGGGGTGGGCCACCCCTTGAAGGGCATTTGTTTACAGATACA 2376
DB 61 ATCCGATCAATGATGGGGGTGGGCCACCCCTTGAAGGGCATTTGTTTACAGATACA 120
QY 2377 TTCTGCAATGATGGAAGAAATGTTTCCGGCCAGAGAGTACACAGACTGGGATACC 2436
DB 121 TTCTGCAATGATGGAAGAAATGTTTCCGGCCAGAGAGTACACAGACTGGGATACC 180
QY 2437 ACTGTGGGCAATGGCAATGTTGATACCTGACACACCTTGTGATGTTATTAATTTA 2496
DB 181 ACTGTGGGCAATGGCAATGTTGATACCTGACACACCTTGTGATGTTATTAATTTA 240
QY 2497 GCAGTTGTGTTATCCGATTAAGAAAGATTAAGGCAAGATTAATGTTGAAGTGCAC 2556
DB 241 GCAGTTGTGTTATCCGATTAAGAAAGATTAAGGCAAGATTAATGTTGAAGTGCAC 300
QY 2557 GCATCTGAAGTCAAACTCTGAGAAAGCTGAATTTGAAGAAATGTTGAATTTATAG 2616
DB 301 GCATCTGAAGTCAAACTCTGAGAAAGCTGAATTTGAAGAAATGTTGAATTTATAG 360
QY 2617 CAAGTGTATTTAGCAACAACCATCTCATCTTAATTTCTTTCACTAACGTGCATTA 2676
DB 361 CAAGTGTATTTAGCAACAACCATCTCATCTTAATTTCTTTCACTAACGTGCATTA 420
QY 2677 A--TTTTTAAAGATATCCCTCTGTGCC--TTTAAATATTCCTAAATTTCTTTT 2733
DB 421 AATTTTAAAGATATTCCTCTGTGCCCTTTTAAATATTCCTAAATTTCTTTT 480
QY 2734 GAGGTGAGT 2743
DB 481 GAGGTGAGT 490
RESULT 14
ABK44858
ID ABK44858 standard; cDNA; 516 BP.
XX
XX ABK44858;
AC
XX
XX
DT 05-JUN-2002 (first entry)
XX
XX
DE cDNA encoding colon tumour protein, SEQ ID NO 409.
XX
XX
XX Human; colon tumour; vaccine; colon cancer; immunogenic;
KM immunotherapy; gene; ss.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200212328-A2.
PN
XX
XX 14-FEB-2002.
PD
XX
XX 31-JUL-2001; 2001WO-US24218.
PF
XX
XX 03-AUG-2000; 2000US-223283P.
PR
XX 28-MAR-2001; 2001US-279763P.
PR
XX 29-JUN-2001; 2001US-302051P.
PR
XX (CORI-) CORIXA CORP.
PA
XX
XX
PI King GE, Meagher MJ, Xu J, Secrist H;
XX
XX WPI; 2002-241739/29.
DR
XX
XX New colon cancer polypeptides and polynucleotides, useful as vaccines,
PT

PT for diagnosing, preventing, and treating colon cancer, and as markers
 for the progression of cancer -

PS Claim 1; SEQ ID No 409; 147bp; English.

XX The invention relates to polynucleotides encoding colon tumour proteins.
 CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
 CC compositions, such as vaccines, for the diagnosis, prevention, and
 CC treatment of colon cancer. Polynucleotide sequences may be used as
 CC hybridisation probes or primers, and in the design and preparation of
 CC ribozyme molecules for inhibiting expression of tumour polypeptides and
 CC proteins in tumour cells. The compositions are useful for stimulating an
 CC immune response against cancer, particularly for the immunotherapy of
 CC colon cancer, and as markers for the progression of cancer.
 CC ABR4450-ABR46237 represent coding sequences of human colon tumour
 CC proteins of the invention.
 CC Note: With the exception of SEQ ID No 1 and 2, the sequence data
 CC for this patent did not form part of the printed specification but was
 CC supplied by the European Patent Office.

XX Sequence 516 BP; 166 A; 112 C; 113 G; 125 T; 0 other;

Query Match 12.9%; Score 431.4; DB 24; Length 516;
 Best Local Similarity 99.8%; Pred. No. 6e-105;
 Matches 432; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 570 CCAGCTTTATTCAGATGTCATCCAGCTTCCCATGATCAACATGTCATCTTTCA 629

DB 1 CCAGCTTTATTCAGATGTCATCCAGCTTCCCATGATCAACATGTCATCTTTCA 60

QY 630 GATCAACACAAACGGAGGAGCATCTCTTACCGAGAGGATCTCAGAAATGAATCC 689

DB 61 GATCAACACAAACGGAGGAGCATCTCTTACCGAGAGGATCTCAGAAATGAATCC 120

QY 690 TGCTAGATCTCTCTATTAATCTGTGATCTCAGTGAAGACATGGAGGCCAGATGA 749

DB 121 TGCTAGATCTCTCTATTAATCTGTGATCTCAGTGAAGACATGGAGGCCAGATGA 180

QY 750 GAATTCCTTCAGTATACCAATCTGTGATATCATAGTGCAGAGAAATTTTGAAGC 809

DB 181 GAATTCCTTCAGTATACCAATCTGTGATATCATAGTGCAGAGAAATTTTGAAGC 240

QY 810 ACCAAACCTGTGAGATGTGAGAAACTCAACTGATCTCAACCCATCAAAATCAGTCA 869

DB 241 ACCAAACCTGTGAGATGTGAGAAACTCAACTGATCTCAACCCATCAAAATCAGTCA 300

QY 870 GGTGCGGTGAATGATCCGGGTGACAAATATCTTGTGACAAAGAGAGTGGCAAG 929

DB 301 GGTGCGGTGAATGATCCGGGTGACAAATATCTTGTGACAAAGAGAGTGGCAAG 360

QY 930 ATTCCATTTTCATATGACCAAGAGATTTTAAGTGAAGTCAAGCCCTTGGACCGA 989

DB 361 ATTCCATTTTCATATGACCAAGAGATTTTAAGTGAAGTCAAGCCCTTGGACCGA 420

QY 990 AGAAAGATGCA 1002

DB 421 AGAAAGATGCA 433

RESULT 15

ABK45154 ID ABR45154 standard; cDNA; 568 BP.

XX ABR45154;

XX 05-JUN-2002 (first entry)

DE cDNA encoding colon tumour protein, SEQ ID No 705.

XX Human; colon tumour; vaccine; colon cancer; immunogenic;

XX immunotherapy; gene; ss.

XX Homo sapiens.

XX PN W0200212328-A2.

XX PD 14-FEB-2002.

XX PF 31-JUL-2001; 2001WO-US24218.

XX PR 03-AUG-2000; 2000US-223283P.

XX PR 28-MAR-2001; 2001US-279763P.

XX PR 29-JUN-2001; 2001US-302051P.

XX PA (CORI-) CORIXA CORP.

XX PI King GE, Meagher MJ, Xu J, Secret H;

XX DR WPI; 2002-241739/29.

XX PT New colon cancer polypeptides and polynucleotides, useful as vaccines,

XX PT for diagnosing, preventing, and treating colon cancer, and as markers

XX PT for the progression of cancer -

XX PS Claim 1; SEQ ID No 705; 147bp; English.

XX The invention relates to polynucleotides encoding colon tumour proteins.
 CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
 CC compositions, such as vaccines, for the diagnosis, prevention, and
 CC treatment of colon cancer. Polynucleotide sequences may be used as
 CC hybridisation probes or primers, and in the design and preparation of
 CC ribozyme molecules for inhibiting expression of tumour polypeptides and
 CC proteins in tumour cells. The compositions are useful for stimulating an
 CC immune response against cancer, particularly for the immunotherapy of
 CC colon cancer, and as markers for the progression of cancer.
 CC ABR4450-ABR46237 represent coding sequences of human colon tumour
 CC proteins of the invention.
 CC Note: With the exception of SEQ ID No 1 and 2, the sequence data
 CC for this patent did not form part of the printed specification but was
 CC supplied by the European Patent Office.

XX Sequence 568 BP; 159 A; 125 C; 94 G; 190 T; 0 other;

Query Match 12.3%; Score 410.4; DB 24; Length 568;
 Best Local Similarity 99.8%; Pred. No. 2.6e-99;
 Matches 411; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2922 CCAGGCTGCTGATCTGACCTGCAAGTGAATGCTGCGCTGGTCCCAATPACAG 2981

DB 1 CCAGGCTGCTGATCTGACCTGCAAGTGAATGCTGCGCTGGTCCCAATPACAG 60

QY 2982 GCATGAACCACTGACCACTTATGATATTTCAATGCTATAGACATTAGAGATTT 3041

DB 61 GCATGAACCACTGACCACTTATGATATTTCAATGCTATAGACATTAGAGATTT 120

QY 3042 TTTGATTTTTCATGACATTTTCTCTCTGCAATGCTTACCTACTTGTGTTTCCC 3101

DB 121 TTTGATTTTTCATGACATTTTCTCTCTGCAATGCTTACCTACTTGTGTTTCCC 180

QY 3102 TTTTGGGGCAAGACAGCATTAATATTTCTGACATTTTCTTATCAGAGATAT 3161

DB 181 TTTTGGGGCAAGACAGCATTAATATTTCTGACATTTTCTTATCAGAGATAT 240

QY 3162 ATCAGTGTGCTCATAGAACCTGCTGATTCATTTATGTTTCTGATTCATCTG 3221

DB 241 ATCAGTGTGCTCATAGAACCTGCTGATTCATTTATGTTTCTGATTCATCTG 300

QY 3222 TGTCCCTTCATCTGACTCTTTGGTATTTCACTGAATTTCAACATTTGTCAAGAA 3281

DB 301 TGTCCCTTCATCTGACTCTTTGGTATTTCACTGAATTTCAACATTTGTCAAGAA 360

QY 3282 GAAAAAGTGAAGCTCAGAAAAATTAATTAATTAAGAACAGCCTTTTGC 3333

DB 361 GAAAAAGTGAAGCTCAGAAAAATTAATTAATTAAGAACAGCCTTTTGC 412

Search completed: June 21, 2003, 21:24:22
Job time : 705 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2003, 21:10:19 ; Search time 183 Seconds
(without alignments)
5605.649 Million cell updates/sec

Title: US-10-025-380-1076

Perfect score: 3345
Sequence: 1 gaattccgctcgaccactg.....cctttcgccgcgaattc 3345

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfil1seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2499	74.7	2499	1	US-08-431-560-2
2	2499	74.7	2499	1	US-08-431-560-2
3	511.4	15.3	627	4	US-09-385-982-514
4	331.4	8.9	346	4	US-09-221-298-21
5	284.8	8.5	591	4	US-09-385-982-300
6	191.8	5.7	3674	4	US-09-301-665-3
7	191.2	5.7	162450	4	US-09-345-882-1
8	191	5.7	4853	4	US-08-881-450A-22
9	190.8	5.7	14581	4	US-08-520-373D-4
10	190.8	5.7	22481	4	US-08-367-841A-43
11	190.8	5.7	22481	5	PCT-US95-07201-43
12	190.8	5.7	22484	4	US-09-875-223-2
13	190.6	5.7	98844	4	US-09-791-211-10
14	190.4	5.7	8453	4	US-09-167-681-45
15	190.4	5.7	36651	4	US-09-738-894A-3
16	189.6	5.7	21234	4	US-09-810-671-3
17	189.6	5.6	70000	4	US-09-851-896-3
18	188	5.6	42571	4	US-09-810-347-3
19	188	5.6	111282	4	US-09-754-250-3
20	186.6	5.6	62804	4	US-09-800-960-3
21	186.4	5.6	956	4	US-09-641-638-56
22	186.4	5.6	20674	4	US-09-641-638-651
23	186.4	5.6	31571	1	US-08-323-443B-1
24	186.4	5.6	53526	3	US-08-658-136-2
25	186.4	5.6	53577	3	US-08-658-136-1
26	186.4	5.6	62804	4	US-09-800-960-3
27	186	5.6	3627	4	US-09-323-873A-6

28	185.6	5.5	631	4	US-09-385-982-354	Sequence 354, App
29	185.4	5.5	87350	3	US-08-781-891-79	Sequence 79, App1
30	185.4	5.5	87543	4	US-09-791-211-3	Sequence 3, App1
31	185	5.5	7676	1	US-08-451-777A-7	Sequence 7, App1
32	185	5.5	7676	2	US-08-451-778A-7	Sequence 7, App1
33	185	5.5	7676	5	PCT-US95-06743-7	Sequence 7, App1
34	185	5.5	7676	2	US-08-998-208-7	Sequence 7, App1
35	184.8	5.5	36741	4	US-09-301-665-3	Sequence 3, App1
36	184.4	5.5	12047	2	US-09-022-461-1	Sequence 3, App1
37	184.4	5.5	12047	4	US-09-033-556-3	Sequence 3, App1
38	184.4	5.5	99500	4	US-09-798-096-10	Sequence 10, App1
39	184.2	5.5	43950	4	US-09-735-934A-3	Sequence 3, App1
40	184.2	5.5	55827	4	US-09-813-133A-3	Sequence 3, App1
41	183.6	5.5	1001	4	US-09-641-638-406	Sequence 406, App
42	183.2	5.5	72604	4	US-09-268-992-7	Sequence 7, App1
43	183.2	5.5	72604	4	US-09-657-474-7	Sequence 7, App1
44	182.4	5.5	6235	4	US-09-305-384-5	Sequence 5, App1
45	182.4	5.5	6679	4	US-09-305-384-1	Sequence 1, App1

ALIGNMENTS

RESULT 1
US-08-431-560-2
; Sequence 2, Application US/08411560
; Patent No. 5620855
; GENERAL INFORMATION:
; APPLICANT: Anne H. Danczig, et al.
; TITLE OF INVENTION: Mammalian Intiflux Peptide
; NUMBER OF INVENTION: Transporter
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,560
; FILING DATE: 01-May-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/013,462
; FILING DATE: 04-FEB-1993
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2499 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-431-560-2
Query Match 74.7%; Score 2499; DB 1; Length 2499;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 88 ATGACTTCTAGCCCATCTTCTCTCTGTCCTTTTATGCTTTATTTGGCAACTGGA 147
DB 1 ATGACTTCTAGCCCATCTTCTCTCTGTCCTTTTATGCTTTATTTGGCAACTGGA 60
QY 148 TATGGCCAAAGAGGAGTTAGTGAGACCCCGAAGCAACCAATTTCTATTTATGAA 207
DB 61 TATGGCCAAAGAGGAGTTAGTGAGACCCCGAAGCAACCAATTTCTATTTATGAA 120
QY 208 GGCCAAAGAACGATCAATATATTTCAAGTTAAGCCAAATCTCTGCTGTGACTTTT 267

121 GGGCAAGAACCGATCAATTAATTCAGTTTAAGGCCAATCTCTGCTGACTTTT 180
122 GAACTAATCTGGGAGACAGACAAATATTTGTATAGAACGGAGGAGCTTGTATTAAC 327
123 GAACTAATCTGGGAGACAGACAAATATTTGTATAGAACGGAGGAGCTTGTATTAAC 240
124 AACAGACCTTGGACAGAGGAAACAAATCTAATCTCAATCTCCAGGTTGACCCCTGGAC 387
125 AACAGACCTTGGACAGAGGAAACAAATCTAATCTCAATCTCCAGGTTGACCCCTGGAC 300
126 GCTAATGGAATTAATAGTGAAGGCTCAGTCCCTATCAACATAGAAAGTGAAGCATCAAC 447
127 GCTAATGGAATTAATAGTGAAGGCTCAGTCCCTATCAACATAGAAAGTGAAGCATCAAC 360
128 GACAAATGACCCACGTTTCTCCAGTCAAAAGTACGAAGGCTCAGTAAGGCAAACTCTGC 507
129 GACAAATGACCCACGTTTCTCCAGTCAAAAGTACGAAGGCTCAGTAAGGCAAACTCTGC 420
130 CAGAGAAAGCCCTTCTGTATGTCAATGCCACAGACTGTGATCCGGCCACTCCCAAT 567
131 CAGAGAAAGCCCTTCTGTATGTCAATGCCACAGACTGTGATCCGGCCACTCCCAAT 480
132 GGGCAGCTTATTAACAGATTTGATCAGCTCAGTCCATGATCAACATGTCATGTACTTT 627
133 GGGCAGCTTATTAACAGATTTGATCAGCTCAGTCCATGATCAACATGTCATGTACTTT 540
134 CAGATCAACAACAAACGGAGCCATCTCTTAACCGAGAGGAGTCTCAGGAATTAAT 687
135 CAGATCAACAACAAACGGAGCCATCTCTTAACCGAGAGGAGTCTCAGGAATTAAT 600
136 CAGATCAACAACAAACGGAGCCATCTCTTAACCGAGAGGAGTCTCAGGAATTAAT 747
137 CAGATCAACAACAAACGGAGCCATCTCTTAACCGAGAGGAGTCTCAGGAATTAAT 660
138 CAGATCAACAACAAACGGAGCCATCTCTTAACCGAGAGGAGTCTCAGGAATTAAT 807
139 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 870
140 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 720
141 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 867
142 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 780
143 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 927
144 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 840
145 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 987
146 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 900
147 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1047
148 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 960
149 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1107
150 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1020
151 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1167
152 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1080
153 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1140
154 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1227
155 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1287
156 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1200
157 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1347
158 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1260
159 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1407
160 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1320
161 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1467
162 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1380
163 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1527
164 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1440
165 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1587
166 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1500
167 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1647
168 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1560
169 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1620
170 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1767
171 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1680
172 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1827
173 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1740
174 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1800
175 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1947
176 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1860
177 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 2007
178 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1920
179 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 2067
180 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 1980
181 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 2127
182 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 2040
183 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 2187
184 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 2100
185 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 2247
186 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 2160
187 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 2307
188 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 2220
189 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 2367
190 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 2280
191 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 2427
192 GAGAAATCTCTCAGTGAATACCAATCTGTGATATCAATAGAGACAGAAATTTTGA 2340

QY 2428 GGGATACCACTGTGGGCGATGGCAGTTGTATACCTGCTGACCAACCTTCTGTGATTTGGT 2487
DB 2341 GGGATACCACTGTGGGCGATGGCAGTTGTATACCTGCTGACCAACCTTCTGTGATTTGGT 2400
QY 2488 ATATATTTTACGAGTGTGTTTATTCGCAATAAAGAGATTAAGCAAAAGTATTTGTA 2547
DB 2401 ATATATTTTACGAGTGTGTTTATTCGCAATAAAGAGATTAAGCAAAAGTATTTGTA 2460
QY 2548 AGTGTCAAGCATCTGAAATGCAAACTCTGAGGAAGCTGA 2586
DB 2461 AGTGTCAAGCATCTGAAATGCAAACTCTGAGGAAGCTGA 2499

RESULT 2
US-08-463-345-2
Sequence 2, Application US/08463345
Patent No. 5710018
GENERAL INFORMATION:
APPLICANT: Anne H. Dantzig, et al.
TITLE OF INVENTION: Mammalian Influx Peptide
TITLE OF INVENTION: Transporter
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN
COUNTRY: U.S.A.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,345
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/013,462
FILING DATE: 04-FEB-1993
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2499 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-463-345-2

Query Match 74.7%; Score 2499; DB 1; Length 2499;
Best Local Similarity 100.0%; Pred.No.0;
Matches 2499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ATGATATCTTCAGGCGCCATCTTCACTCCCTGTGTCTTATGCTTATTTGGCACTGGA 147
DB 1 ATGATATCTTCAGGCGCCATCTTCACTCCCTGTGTCTTATGCTTATTTGGCACTGGA 60
QY 148 TATGCGCAAGAGGGGAAAGTTAGTGAAGCCCTGAAAACCATGACATTTTCTATTATGA 207
DB 61 TATGCGCAAGAGGGGAAAGTTAGTGAAGCCCTGAAAACCATGACATTTTCTATTATGA 120
QY 208 GGGCAAGAACCGGATGAATTAATTCAGGTTTAAAGCCCATCTCTCTGTGACTTTT 267
DB 121 GGGCAAGAACCGGATGAATTAATTCAGGTTTAAAGCCCATCTCTCTGTGACTTTT 180
QY 268 GAATTAAGTGGGAGACAGACATATTTGTATAGAAAGGAGGAGCTTCTGATTAAC 327
DB 181 GAATTAAGTGGGAGACAGACATATTTGTATAGAAAGGAGGAGCTTCTGATTAAC 240
QY 328 AACAGAGCTTGGACAGGGGAAACAGATCTACTCAATCTCCAGTTGACAGCCTTGAC 387

DB 241 AACAGAGCTTGGACAGGGAAACAGATCTACTCAATCTCCAGTTGACAGCCTTGAC 300
QY 388 GCTAATGAAATTAATGAGGGGTCCAGTCCCATGACCATGAAATGAGACATCAAC 447
DB 301 GCTAATGAAATTAATGAGGGGTCCAGTCCCATGACCATGAAATGAGACATCAAC 360
QY 448 GACAAATGACCAAGCTTCTCCAGTCAAGTGAAGAGGTCTAGTAAGGACAGACTCCG 507
DB 361 GACAAATGACCAAGCTTCTCCAGTCAAGTGAAGAGGTCTAGTAAGGACAGACTCCG 420
QY 508 CCAAGAAAGCCTTCTGTTATGATCAATGCAAGACACTGATGATCCGCACTCCCAAT 567
DB 421 CCAAGAAAGCCTTCTGTTATGATCAATGCAAGACACTGATGATCCGCACTCCCAAT 480
QY 568 GGCAGAGCTTAATGACAGATGTCATCCAGTCCCATGATCAAGATGATGATCTT 627
DB 481 GGCAGAGCTTAATGACAGATGTCATCCAGTCCCATGATCAAGATGATGATCTT 540
QY 628 CAGATCAACAAACAAACGGAGCCATCTCTTAAACCGAGAGGATCTGAGAAATTAAT 687
DB 541 CAGATCAACAAACAAACGGAGCCATCTCTTAAACCGAGAGGATCTGAGAAATTAAT 600
QY 688 CCGCTTAAGAAATCTCTTAATATCTGATGATCTCAAGTGAAGACATGGAGCCAGAT 747
DB 601 CCGCTTAAGAAATCTCTTAATATCTGATGATCTCAAGTGAAGACATGGAGCCAGAT 660
QY 748 GGAATTCCTCAGGATGACATCTGATGATGATGATGATGATGATGATGATGATGAT 807
DB 661 GGAATTCCTCAGGATGACATCTGATGATGATGATGATGATGATGATGATGATGAT 720
QY 808 GCACCAAAACCTGTGGAGATGGTGGAAAACCTCAACTGATCTCAACCCATCAAAATCACT 867
DB 721 GCACCAAAACCTGTGGAGATGGTGGAAAACCTCAACTGATCTCAACCCATCAAAATCACT 780
QY 868 CAGGTGCGGTGAATGATCCGCTGACAAATATCTTATGTTGAACAAAGAAAGCTGCCA 927
DB 781 CAGGTGCGGTGAATGATCCGCTGACAAATATCTTATGTTGAACAAAGAAAGCTGCCA 840
QY 928 AGATTCCTCAATTTCAATTTGACAGAGAGATTTTACGTGATCCAGCCCTTGGACCA 987
DB 841 AGATTCCTCAATTTCAATTTGACAGAGAGATTTTACGTGATCCAGCCCTTGGACCA 900
QY 988 GAGAAAGAGATGATATGTTTATGAGTGGCAAGAGATGATGATGATGATGATGATGAT 1047
DB 901 GAGAAAGAGATGATATGTTTATGAGTGGCAAGAGATGATGATGATGATGATGATGAT 960
QY 1048 TCATATCCGCTGAAATTCATGTAAGTTAAAGATTAATGATGATGATGATGATGAT 1107
DB 961 TCATATCCGCTGAAATTCATGTAAGTTAAAGATTAATGATGATGATGATGATGATGAT 1020
QY 1108 CCGTCAACAGTAACCGTATTTTGAAGTCCAGAGAAATGAAACGATGGTAAACGATTCGG 1167
DB 1021 CCGTCAACAGTAACCGTATTTTGAAGTCCAGAGAAATGAAACGATGGTAAACGATTCGG 1080
QY 1168 ACCCTTACGACATGACAGAGATGAAGAAATATCTGCCAAGTCTTCTTAATCAACAG 1227
DB 1081 ACCCTTACGACATGACAGAGATGAAGAAATATCTGCCAAGTCTTCTTAATCAACAG 1140
QY 1228 ATTTGAGAGCAAACTCCCAACTTCCATGATGATGATGATGATGATGATGATGATGAT 1287
DB 1141 ATTTGAGAGCAAACTCCCAACTTCCATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1288 GGAATGTTACAGTTAGCTTAACAGTCTTGAAGAAAGCAAGATCTCTCAAGTAACTTA 1347
DB 1201 GGAATGTTACAGTTAGCTTAACAGTCTTGAAGAAAGCAAGATCTCTCAAGTAACTTA 1260
QY 1348 ACGATAGAGGTGTGACAAAGATTTCAAGACCTTGTGTTGTCGAATCAAGCTTAAT 1407
DB 1261 ACGATAGAGGTGTGACAAAGATTTCAAGACCTTGTGTTGTCGAATCAAGCTTAAT 1320
QY 1408 GATATCAATGATCAGATCCCATCTTTGAAAAATCAGATTATGAAACCTGACTTGTGT 1467
DB 1321 GATATCAATGATCAGATCCCATCTTTGAAAAATCAGATTATGAAACCTGACTTGTGT 1380

QY 1468 GAAGACAAAATTGGGTCACCATTTTAAACCTCCAGGCCACTGATGCTGATGAGCCA 1527
DB 1381 GAAGACAAAATTGGGTCACCATTTTAAACCTCCAGGCCACTGATGCTGATGAGCCA 1440
QY 1528 TTTACTGGGAGTCTAAATTTCTGTATCATATCATTAAGGAGACAGTGAAGGAGCCCTG 1587
DB 1441 TTTACTGGGAGTCTAAATTTCTGTATCATATCATTAAGGAGACAGTGAAGGAGCCCTG 1500
QY 1588 GGGGTTGACACAGATCCCATACCAACACCGGATATGTCATTAATAAAAGCCTCTTGAT 1647
DB 1501 GGGGTTGACACAGATCCCATACCAACACCGGATATGTCATTAATAAAAGCCTCTTGAT 1560
QY 1648 TTTGAAACAGCAGCTGTTTCCACATTTGTTCTTAAGACAAAATCCTGAGCCTCTAGT 1707
DB 1561 TTTGAAACAGCAGCTGTTTCCACATTTGTTCTTAAGACAAAATCCTGAGCCTCTAGT 1620
QY 1708 TTTGGTGTGAAGTACATAGCAAGTTCTTTGGCCAAAGTCCAGCTTATTTGACAGATGTG 1767
DB 1621 TTTGGTGTGAAGTACATAGCAAGTTCTTTGGCCAAAGTCCAGCTTATTTGACAGATGTG 1680
QY 1768 AATGAAGCACTCAATTTTCCCAACAGTATTCACAGCGAAAGTCACTGAGATGACT 1827
DB 1681 AATGAAGCACTCAATTTTCCCAACAGTATTCACAGCGAAAGTCACTGAGATGACT 1740
QY 1828 ATAGGCACTAAAGTGGGCAATGTGACTGCCAAGATCCAAAGGTCCTGACATTAAGCTAT 1887
DB 1741 ATAGGCACTAAAGTGGGCAATGTGACTGCCAAGATCCAAAGGTCCTGACATTAAGCTAT 1800
QY 1888 TCACTGAGGGGAGACACAAGAGTGTGCTTAAATTAATGACACGTCGTGATGATCTTT 1947
DB 1801 TCACTGAGGGGAGACACAAGAGTGTGCTTAAATTAATGACACGTCGTGATGATCTTT 1860
QY 1948 AGTGTGCTCCTCATTGACAGAGAAAGCCGAAAGTCCATATGCGGTGCAAGTGTGCGACA 2007
DB 1861 AGTGTGCTCCTCATTGACAGAGAAAGCCGAAAGTCCATATGCGGTGCAAGTGTGCGACA 1920
QY 2008 GAAATGAGGGGGGTCTTCTTAAAGTCTGTGTCAGAGTTTCCACTGATCTTAAGGATGTG 2067
DB 1921 GAAATGAGGGGGGTCTTCTTAAAGTCTGTGTCAGAGTTTCCACTGATCTTAAGGATGTG 1980
QY 2068 AATGACAAACCTCCAGGCTGACCAAGGACTTACAGAGGAGTCTTCTGCGATCCCTC 2127
DB 1981 AATGACAAACCTCCAGGCTGACCAAGGACTTACAGAGGAGTCTTCTGCGATCCCTC 2040
QY 2128 AGTGCACCTGGAAAGTCTCAATTTTGCAGGCTACTGATGATGATGACACTTAATTTGCGGGT 2187
DB 2041 AGTGCACCTGGAAAGTCTCAATTTTGCAGGCTACTGATGATGATGACACTTAATTTGCGGGT 2100
QY 2188 CCCCATTTTACATTTTCCCTGCGAGTGGAGGCTTACAAAACGACGAGGAGTTTCCAAA 2247
DB 2101 CCCCATTTTACATTTTCCCTGCGAGTGGAGGCTTACAAAACGACGAGGAGTTTCCAAA 2160
QY 2248 ATCAATGTACTCATGCGGAGCTGTCTACAGGACACAGACTTTGAGAGAGGCGCTAT 2307
DB 2161 ATCAATGTACTCATGCGGAGCTGTCTACAGGACACAGACTTTGAGAGAGGCGCTAT 2220
QY 2308 GTGCTCTTGATCCGATCAATGATGAGGGGTGCGCCACCTTGAAGGCAATTTCTTTA 2367
DB 2221 GTGCTCTTGATCCGATCAATGATGAGGGGTGCGCCACCTTGAAGGCAATTTCTTTA 2280
QY 2368 CCAATTCATTTCTGCAATTTGTTGGAAGAGTGTTCGCGGCAAGAGCTGACACACT 2427
DB 2281 CCAATTCATTTCTGCAATTTGTTGGAAGAGTGTTCGCGGCAAGAGCTGACACACT 2340
QY 2428 GGAATACCACTGTGGGCAATGGAGTGTATACGTCGACACACCTTCTGATTTGGT 2487
DB 2341 GGAATACCACTGTGGGCAATGGAGTGTATACGTCGACACACCTTCTGATTTGGT 2400
QY 2488 ATAAATTTAGCAGTTGTGTTTATCCGATTAAGAGAGATTAAGGCAAGATTAATGTGAA 2547
DB 2401 ATAAATTTAGCAGTTGTGTTTATCCGATTAAGAGAGATTAAGGCAAGATTAATGTGAA 2460

QY 2548 AGTGCACAGCATCTGAAGTCAAACTCTGAGAGCTGA 2586
DB 2461 AGTGCACAGCATCTGAAGTCAAACTCTGAGAGCTGA 2499

RESULT 3
US-09-385-982-514
; Sequence 514, Application US/09385982
; Patent No. 6262314
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCNDA-260XX
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 514
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(627)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-514

Query Match 15.3%; Score 511.4; DB 4; Length 627;
Best Local Similarity 99.0%; Pred. No. 1e-132;
Matches 512; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2254 GGTACTCAGTCCCGGAGTGTCTACAGGACACAGACTTTGAGGAGGGGCTATGTCGTC 2213
DB 1 GGTACTCAGTCCCGGAGTGTCTACAGGACACAGACTTTGAGGAGGGGCTATGTCGTC 60
QY 2314 TTGATCCGATCAATATGAGGGGTGCGGCAACCTTGAAGGCAATTTCTTTACAGATT 2373
DB 61 TTGATCCGATCAATATGAGGGGTGCGGCAACCTTGAAGGCAATTTCTTTACAGATT 120
QY 2374 ACATTTGCAATTTGTTGGAAGGAGTGTTCCTGCGGACAGAGCTACCCAGACTGGGATA 2433
DB 121 ACATTTGCAATTTGTTGGAAGGAGTGTTCCTGCGGACAGAGCTACCCAGACTGGGATA 180
QY 2434 CCCACTGTGGGCAATGGCAAGTGTGATACGCTGACCAACCTTGTGGATTTGATTAAT 2493
DB 181 CCCACTGTGGGCAATGGCAAGTGTGATACGCTGACCAACCTTGTGGATTTGATTAAT 240
QY 2494 TTGACAGTTGTGTTTATCCGATTAAGAGATTAAGGCAAGATTAATGTTGAAGTCT 2553
DB 241 TTGACAGTTGTGTTTATCCGATTAAGAGATTAAGGCAAGATTAATGTTGAAGTCT 300
QY 2554 CAAGCATCTGAAGTCAAACTCTGAGAGCTGAATTTGAAAAGGATGTTGAATTTATA 2613
DB 301 CAAGCATCTGAAGTCAAACTCTGAGAGCTGAATTTGAAAAGGATGTTGAATTTATA 360
QY 2614 TAGCAAGTGTATTTAGCAACAACATGCTCATCTTAATCTTATCATCTTAAGTGCATT 2673
DB 361 TAGCAAGTGTATTTAGCAACAACATGCTCATCTTAATCTTATCATCTTAAGTGCATT 420
QY 2674 ATAAATTTTAAACAGATATTCCTCTGCTGCTTAATATTTGCTAAATATTTCTTTT 2733
DB 421 ATAAATTTTAAACAGATATTCCTCTGCTGCTTAATATTTGCTAAATATTTCTTTT 480
QY 2734 GAGGTGAGTCTGCTCTGTCGCCAGGCTGAGAGTAC 2770
DB 481 GAGGTGAGTCTGCTCTGTCGCCAGGCTGAGAGTAC 517

```
RESULT 4
US-09-221-298-21/c
; Sequence 21, Application US/09221298
; Patent No. 6284241
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER
; FILE REFERENCE: 210121.471
; CURRENT APPLICATION NUMBER: US/09/221,298
; CURRENT FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 346
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (257)
; OTHER INFORMATION: where n is a, c, g or t
US-09-221-298-21

Query Match
Best Local Similarity 9.9%; Score 331.4; DB 4; Length 346;
Matches 343; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1661 CTGTTTCCACATTGTGTTCAGAGCAAAATCTGAGCTCTGAGTGTGTGTAAGT 1720
DB 346 CTGTTTCCACATTGTGTTCAGAGCAAAATCTGAGCTCTGAGTGTGTGTAAGT 287
QY 1721 ACAATGCAAGTCTTTTGGCAAGTTACG-CTTATTGTGACAGTGTGAAGCACT 1779
DB 286 ACAATGCAAGTCTTTTGGCAAGTTACG-CTTATTGTGACAGTGTGAAGCACT 227
QY 1780 CAATTTTCCACACGATTTTCCAGAGCAAAATGAGTGTGAGTGTGAGTGTGAGT 1839
DB 226 CAATTTTCCACACGATTTTCCAGAGCAAAATGAGTGTGAGTGTGAGTGTGAGT 167
QY 1840 GTGGGCAATGTGACTGCGCAAGATCCAGAGTGTGACATTAAGTGTGAGGGA 1899
DB 166 GTGGGCAATGTGACTGCGCAAGATCCAGAGTGTGACATTAAGTGTGAGGGA 107
QY 1900 GACACAGAGGTTGGCTTAATTTGACAGTGTGAGTGTGAGTGTGAGTGTGAGT 1959
DB 106 GACACAGAGGTTGGCTTAATTTGACAGTGTGAGTGTGAGTGTGAGTGTGAGT 47
QY 1960 TTGACAGAGAGACCCGGAATCCATATCGGTTACAGTGTGAGTGTGAGT 2004
DB 46 TTGACAGAGAGACCCGGAATCCATATCGGTTACAGTGTGAGTGTGAGT 2

RESULT 5
US-09-385-982-300/c
; Sequence 300, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS: II
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 300
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(591)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-300

Query Match
Best Local Similarity 8.5%; Score 284.8; DB 4; Length 591;
Matches 286; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2480 TGATGGTATTAATTTAGAGTGTGTATCCGCAATTAAGAGATTAAGCAAGAT 2539
DB 289 TCAATGGTATTAATTTAGAGTGTGTATCCGCAATTAAGAGATTAAGCAAGAT 230
QY 2540 ATGTTGAAGTGTCAAGATCTGAAGTCAAACTCTGAGAGCTGAATTTGAAGAA 2599
DB 229 ATGTTGAAGTGTCAAGATCTGAAGTCAAACTCTGAGAGCTGAATTTGAAGAA 170
QY 2600 TGTTTGAATTTATATGCAAGTCTATTTTCAGCAAAACCATCTCATCTATTCT 2659
DB 169 TGTTTGAATTTATATGCAAGTCTATTTTCAGCAAAACCATCTCATCTATTCT 110
QY 2660 ATTAAGTGCATTAATTTTAAACAGATATTCCTGCTGCTTAAATTTGCTA 2719
DB 109 ATTTAAGTGCATTAATTTTAAACAGATATTCCTGCTGCTTAAATTTGCTA 50
QY 2720 AATATTTCTTTTGTGAGTGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 2768
DB 49 AATATTTCTTTTGTGAGTGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1
```

```
RESULT 6
US-09-301-665-3
; Sequence 3, Application US/09301665
; Patent No. 6207876
; GENERAL INFORMATION:
; APPLICANT: KELLEMS, RODNEY E.
; APPLICANT: DATTA, SURJIT K.
; TITLE OF INVENTION: ADENOSINE DEAMINASE DEFICIENT TRANSGENIC MICE AND
; TITLE OF INVENTION: METHODS FOR THE USE THEREOF
; FILE REFERENCE: UTS-243
; CURRENT APPLICATION NUMBER: US/09/301,665
; CURRENT FILING DATE: 1999-04-28
; EARLIER APPLICATION NUMBER: 60/083,408
; EARLIER FILING DATE: 1998-04-29
; EARLIER APPLICATION NUMBER: 60/083,370
; EARLIER FILING DATE: 1998-04-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 36741
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-301-665-3

Query Match
Best Local Similarity 5.7%; Score 191.8; DB 4; Length 36741;
Matches 217; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 2717 CTAAATATTTCTTTTGTGAGTGTGAGTCTTGTGTGCGCCAGGCTGAGTGAAGTGT 2776
DB 19604 CTTTCTTTTCTTTTGTGAGTGTGAGTCTTGTGTGCGCCAGGCTGAGTGAAGTGT 19663
QY 2777 GTGATCCAGTCACATCGCAACTCGGCTCGGCTGCTGCTGCTGCTGCTGCTGCT 2836
DB 19664 ATGATCTCGGCTCAGCACTCCAGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 19723
QY 2837 TCTTAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2896
```



```
Db      19724 TCCGAGTACCTGGATTACAGACATGCCACCACTCAGCTAATTTTGTATTTT 19783
QY      2897 ATAGAGACGGGTTTCCGCAATTTGGCCAGGCTGTCTTGAATCTTCAGCTCAAGTGATC 2956
Db      19784 GCGGAGATGGGGTTTTCATGTTGGCCAGGCTGTGTCTGAACTCTTACCTCAAGTATC 19843
QY      2957 TGCCTGCTGTGTCTTCCCA 2975
Db      19844 CACCCAGCTTGGCTCCCA 19862

RESULT 7
US-09-345-882-1
/ Sequence 1, Application US/09345882
/ Patent No. 6393373
/ GENERAL INFORMATION:
/ APPLICANT: Bouguenere, Lydie
/ TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
/ FILE REFERENCE: GENSET.031A
/ CURRENT APPLICATION NUMBER: US/09/345,882
/ PRIOR FILING DATE: 1999-06-30
/ PRIOR APPLICATION NUMBER: US 60/091,315
/ PRIOR FILING DATE: 1998-06-30
/ PRIOR APPLICATION NUMBER: US 60/111,909
/ PRIOR FILING DATE: 1998-12-10
/ NUMBER OF SEQ ID NOS: 140
/ SOFTWARE: Patent.pm
/ SEQ ID NO 1
/ LENGTH: 162450
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 72794
/ OTHER INFORMATION: 5-124-273 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 88073
/ OTHER INFORMATION: 5-127-261 : polymorphic base A or C
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 90842
/ OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 93714
/ OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 97122
/ OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 97152
/ OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 99098
/ OTHER INFORMATION: 5-130-257 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 99117
/ OTHER INFORMATION: 5-130-276 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 103806
/ OTHER INFORMATION: 5-131-395 : polymorphic base A or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 106940
/ OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A

/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 108106
/ OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 108149
/ OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 108308
/ OTHER INFORMATION: 5-135-357 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 108471
/ OTHER INFORMATION: 5-136-174 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 134134
/ OTHER INFORMATION: 5-140-120 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 134362
/ OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 134374
/ OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 146328
/ OTHER INFORMATION: 5-143-84 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 146345
/ OTHER INFORMATION: 5-143-101 : polymorphic base A or C
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 150329
/ OTHER INFORMATION: 5-145-24 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 160031
/ OTHER INFORMATION: 5-148-352 : polymorphic base G or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 72771..72817
/ OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 72771..72817
/ OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 88050..88096
/ OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 88050..88096
/ OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 90819..90865
/ OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 90819..90865
/ OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 93690..93736
/ OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
/ FEATURE:
```

```

NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:
Query Match 5.7% Score 191.2; DB 4; Length 162450;
Best Local Similarity 86.7%; Pred.No.3.2e-42;
Matches 222; Conservative 0; Mismatches 33; Indels 1; Gaps 1

```

Qy	2722	TATTTCTTTTAAAGGAGAGCTTGCTGCTGTCGCCAGGCTGAGTAACAGTGTGAT	2781
Dd	86094	TTTTTTTTTTTAAAGAGAGTCTCACTCTGTGCCAGGCTGAGTGAAGTGTGCAT	86153
Qy	2782	CCGAGCTCACTGCAACCTCGCTCTCGGAGTTTCACATGATTTCTCTGCTCAAGTCTCTA	2841
Dd	86154	ATCGGCTCACTGCAACCTCGCTCTCGGAGTTTCAGGCAATTTCTCTGCTCAAGCTCCAG	86213
Qy	2842	AGTAGCTGGGTTTACAGGACCCACGACATGCGCAGTAATTTTTGTATTTTAAATAA	2901
Dd	86214	AGTAGCTGGGATTTACAGGACCCGCGCACCATGCTGCTTAATTTTGTATTTTAAATAA	86273
Qy	2902	GACGGGGTTTGGCCA-TTTGGCCAGGCTGCTTGAACCTCTGACGCTCAAGTATCTGCC	2960
Dd	86274	GACGAGGTTTCAACATTTTGGCCAGGCTGCTCTGAACCTCTGACCTCAGTGAATCCACC	86333
Qy	2961	TGCTTGCTGCCCA 2976	
Dd	86334	TGCTTGCTGCCCA 86349	

RESULT 8
US-08-881-450A-22

Sequence 22, Ap

GENERAL INFORMATION

APPLICANT: Habene

TITLE OF INVENTION:	COMPOSITIONS AND METHODS
1. A method for determining the concentration of a substance in a sample, comprising the steps of: (a) measuring the absorbance of the sample at a wavelength of 254 nm; (b) measuring the absorbance of a standard solution of the substance at a wavelength of 254 nm; and (c) comparing the absorbance of the sample to the absorbance of the standard solution.	1. A method for determining the concentration of a substance in a sample, comprising the steps of: (a) measuring the absorbance of the sample at a wavelength of 254 nm; (b) measuring the absorbance of a standard solution of the substance at a wavelength of 254 nm; and (c) comparing the absorbance of the sample to the absorbance of the standard solution.

TITLE OF INVENTION: PANCREATIC DISEASE
NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Wilcoff, Inc

STREET: One
CITY: Boston

STATE: Massa

COUNTRY: USA

ZIP: 02111
COMPUTER READAB

COMPUTER READABLE
MEDIUM TYPE:

COMPUTER: IB

OPERATING SYS

```

; SOFTWARE: WORKING SOFTWARE
; CURRENT APPLICATION:

```

CONSENT AND ELECT
APPLICATION N

FILING DATE:

CLASSIFICATIO
PRIOB APPLICATI

APPLICATION N

FILING DATE:

ATTORNEY/AGENT
NAME: Kathleen

NAME: NALINE
REGISTRATION

REFERENCE/DOC

TELECOMUNICATI
TELEFONIA
;

TELEPHONE: 617
TELEFAX: 617

; INFORMATION FOR S

SEQUENCE CHARAC

```

;
; LENGTH: 4853
; TYPE: miciei

```

STRANDEDNESS:

```

;      TOPOLOGY: 11
;      -----

```

MOLECULE TYPE:
FEATTYPE:
;
:

NAME/KEY: IP

US-08-881-450A-22

Answer Match

Query MATCH
Best Local Similar

Matches 227; Con

2000 2000

2630 MILES

— — —

[illegible]

```

RESULT 9
US-08-520-373D-4
: Sequence 4, Application US/08520373D
: Patent No. 6451763
: GENERAL INFORMATION:
: APPLICANT: Tombrian-Tink, Joyce
: APPLICANT: Steele, Fintan R
: APPLICANT: Chader, Gerald J
: APPLICANT: Becerra, Sofia P
: APPLICANT: Johnson, Lincoln V
: APPLICANT: Rodriguez, Ignacio R
: TITLE OF INVENTION: RETINAL PIGMENTED EPITHELIUM DERIVED NEUROTROPIC FACTOR
: FILE REFERENCE: 2026-4203US1
: CURRENT APPLICATION NUMBER: US/08/520,373D
: CURRENT FILING DATE: 1995-08-29
: PRIOR APPLICATION NUMBER: 08/377,710
: PRIOR FILING DATE: 1995-01-25
: PRIOR APPLICATION NUMBER: 08/779,979
: PRIOR FILING DATE: 1994-07-25
: PRIOR APPLICATION NUMBER: 07/994,215
: PRIOR FILING DATE: 1992-06-04
: PRIOR APPLICATION NUMBER: 07/952,796
: PRIOR FILING DATE: 1992-09-24
: NUMBER OF SEQ ID NOS: 34
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 4
: LENGTH: 14581
: TYPE: DNA
: ORGANISM: HUMAN
: FEATURE:
: OTHER INFORMATION: mRNA: 6683; EXON: 6683-6790; EXON 11584-11675;
: OTHER INFORMATION: EXON: 14539-14581; INTRON: 6791-11583; INTRON:
: OTHER INFORMATION: 11676-14538; CDS: 11584-11675; 14539-14580
: US-08-520-373D-4

```

[illegible]

QY		9782
Db		9782

```

1      RESULT 10
2      US-08-367-841A-43
3      : Sequence 43, Application US/08367841A
4      : Patent No. 6319687
5      : GENERAL INFORMATION:
6      : APPLICANT: Chader, Gerald J.; Rodriguez,
7      : APPLICANT: Ignacio R.; Mazuruk, Kzyzstof;
8      : APPLICANT: Tombran-Tink, Joyce
9      : TITLE OF INVENTION: PIGMENT EPITHELIUM
10     : TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
11     : NUMBER OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
12     : NUMBER OF SEQUENCES: 43
13     : CORRESPONDENCE ADDRESS:
14     : ADDRESSEE: Morgan & Flinnegan
15     : STREET: 345 Park Avenue
16     : CITY: New York
17     : STATE: New York
18     : COUNTRY: USA
19     : ZIP: 10154
20     : COMPUTER READABLE FORM:
21     : MEDIUM TYPE: Floppy disk
22     : COMPUTER: IBM PC Compatible
23     : OPERATING SYSTEM: PC-DOS/MS-DOS
24     : SOFTWARE: WORDPERFECT 5.1
25     : CURRENT APPLICATION DATA:
26     : APPLICATION NUMBER: US/08/367,841A
27     : FILING DATE: 30-DEC-1994
28     : CLASSIFICATION: 435
29     : PRIOR APPLICATION DATA:
30     : APPLICATION NUMBER: 08/257,963
31     : FILING DATE: 07-JUN-1994
32     : PRIOR APPLICATION DATA:
33     : APPLICATION NUMBER: 07/952,796
34     : FILING DATE: 24-SEP-1992
35     : ATTORNEY/AGENT INFORMATION:
36     : NAME: DOROTHY R. AUTH
37     : REGISTRATION NUMBER: 36434
38     : REFERENCE/DOCKET NUMBER: 20264126US2
39     : TELECOMMUNICATION INFORMATION:
40     : TELEPHONE: (212) 759-4800
41     : TELEFAX: (212) 751-6849
42     : INFORMATION FOR SEQ ID NO: 43:
43     : SEQUENCE CHARACTERISTICS:
44     : LENGTH: 22481 Base Pairs
45     : TYPE: Nucleic Acid
46     : STRANDEDNESS: Double
47     : TOPOLOGY: Unknown
48     : MOLECULE TYPE: Genomic DNA
49     : FEATURE:
50     : NAME/KEY: Pl-147
51     : LOCATION:
52     : IDENTIFICATION METHOD:
53     : OTHER INFORMATION: full length genomic
54     : OTHER INFORMATION: sequence for PEDF plus flanking sequences.
55     : US-08-367-841A-43

```

	Query Match	5.7%;	Score 190.8;	DB 4;	Length 22481;
	Beet Local Similarity	82.7%;	Pred. No. 1,4e-42;		
	Matches 230;	Conservative	0;	Mismatches 47;	Indels 1; Gaps 1;
QY	2700 TTGCTTAAATTTTGCTAATATTCTTTTTGAGGTGGAGAGCTGCTGTGGGCCCA				2759
Db	9494 TTCTACATTTTTTTTTTCCTTTTTTTTTTTTTTTTGAATGAGCTGTCTGTGGCCA				9553
QY	2760 GGCTGGAGTACAGTGGGTATCCACAGCTCACTGCACACTCGGCTCCTGGTTCACATG				2819
Db	9554 GCCTGGAGTACAGTGGCGCATCTCACTACTACGCAACTCCACTTCTTGATTTCAAGTG				9613

Db 9674 TAAATTTTGTATTTTGTAGAGATGGGGTTTCCACGATTGGCCAGATGTCCTCAAC 9733

Qy 2939 TCCCTGACGTCAAGTATCTGCCTGCTTGGATCTCCAA 2976

Db 9734 TCCCTGACCTCAGGTGATCTACCCGCTGGCCTCTCA 9771

RESULT 13
US-09-791-211-10

Query Match	5.7%	Score 190.6	DB 4	Length 9884
Best Local Similarity	71.2%	Pred. No. 3.6e-42		
Matches 280		Conservative 0	Mismatches 109	Indels 4
				Gaps 2

Qy	Dy	Qy	Dy
33075	TATGATGAGACGGGGTTTCCCATATGCCAGGCTGGTTCGAAACCCCTGGCCCTCAAAATCA	33134	
2955	TCGCGCTGCCTGGTGTTCCTCAA--TACAGACATGAAACAACCTGACACCCAC--CTACCTAGA	3010	
33135	TCGCGCTGCCTGGCTCTCCCAAGTACTGGGATTTACACACGTAGAGCACTGGCGCTTAAT	33194	
3011	TATTTCATGTGCTATAGACATAGAGAGATTTT	3043	
33195	TTTTTAATTGCTTTTTCACCTGAAATATTAAT	33227	

```

RESULT 14
US-09-167-681-45
; Sequence 45, Application US/09167681A
; Patient No. 6265561
; GENERAL INFORMATION:
; APPLICANT: Welnshillboun, M.D., Richard M.
; APPLICANT: Rathjogianis, Rebecca B.
; APPLICANT: Wood, Thomas C.
; APPLICANT: Ottensmeyer, Diane M.
; TITLE OF INVENTION: SULFOTANSFERASE SEQUENCE VARIANTS
; FILE REFERENCE: 07039/118001
; CURRENT APPLICATION NUMBER: US/09/167,681A
; CURRENT FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 45
; LENGTH: 8447
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4361) ... (4507)
; NAME/KEY: CDS
; LOCATION: (4612) ... (4737)
; NAME/KEY: CDS
; LOCATION: (4827) ... (4925)
; NAME/KEY: CDS
; LOCATION: (6322) ... (6447)
; NAME/KEY: CDS
; LOCATION: (6543) ... (6638)
; NAME/KEY: CDS
; LOCATION: (7137) ... (7316)
; NAME/KEY: CDS
; LOCATION: (7439) ... (7553)
; US-09-167-681-45

```

Query Match	5.7%;	Score 190.4;	DB 4;	Length 8453;
Best Local Similarity	75.3%;	Pred. No. 1,1e-42;		
Matches 250;	Conservative 0;	Mismatches 81;	Indels 1;	Gaps 1;

Db 1930 CCTCAGGTGATCCTGCTGCATCGACCTCCAA 1961

RESULT 15

```

US-09-738-894A-3/C
; Sequence 3, Application US/09738894A
; Patent No. 6331423
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C1000636
; CURRENT APPLICATION NUMBER: US/09/738, 894A
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 36651
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(36651)
; OTHER INFORMATION: n = A,T,C or G
US-09-738-894A-3

```

Query Match	5.7%;	Score 190.4;	DB 4;	Length 36651;
-------------	-------	--------------	-------	---------------

Matches 232; Conservative 0; Mismatches 51; Indels 1; Gaps 1;

Qy	2694	TCCTCTTTCCTTTATATATTCCTAAATTTCTTTTGTAGTGGAGCTGTCTGT	2753
Db	15854	TTCTTCTATACCTCAATTTTTTTTTTTTTTTTTTTTATGATGGAGTCTGTCTGT	15739
Qy	2754	CGCCCAAGCTGAGATACAGTGTGTGTATCCCAAGCTACGTCAAACCTCCGCTCTGGGTT	2813
Db	15794	CTCCCAAGCTGAGATGCAATGGGATGATTTTCGGCTCATCAACAACCTCCGCTTCGGGTT	15739
Qy	2814	CACATGATTTCTCGCTCAGCTTCTCTTAATGATCTGGGTTTACAGGACACCAACACATG	2873
Db	15734	CAGGCAATTTCTCTACTCTCAGCTTCCCAATGATGGGATTTATGGGGCCCGCACCA	15675
Qy	2874	CCAGCTAATTTTTGTATTTTTTAATAGACGGGGTTTGCCAT-TTGGCCAGCTGTGTC	2932
Db	15674	CCAGCTAATTTTTGTATTTTTTGTAGATACAGGTTTCAACATGTTGGCCAGGCTGTGTC	15615
Qy	2933	TTGAATCTCCGAGCTGAAGGATCTGCTGCCCTTGTGTCGCCAA	2976
Db	15614	TGAATCTCTGACTCAAGTATCTGCCACTTGGCTTCCAA	15571

```
Search completed: June 22, 2003, 00:55:41
Job time : 187 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2003, 21:12:29 / Search time 447 Seconds
(without alignments)
10981.083 Million cell updates/sec

Title: US-10-025-380-1076

Perfect score: 3345
Sequence: 1 gaattccgcttcgacacatg.....cctttgcgcgcgaattc 3345

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PTCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3345	100.0	3345	9	US-10-025-380-1076
2	3345	100.0	3345	10	US-09-922-217-1076
3	3345	100.0	3345	10	US-09-862-436-302
4	3345	100.0	3345	10	US-09-833-263-1076
5	3345	100.0	3345	10	US-09-880-107-3288
6	3314.8	99.1	3654	10	US-09-862-832-118
7	3314.8	99.1	3654	10	US-09-880-107-3807
8	2432.6	72.7	2877	9	US-10-025-380-1086
9	2432.6	72.7	2877	10	US-09-922-217-1086
10	2432.6	72.7	2877	10	US-09-833-263-1086
11	514.8	15.4	518	10	US-09-998-598-1266
12	514.8	15.4	520	9	US-10-046-935-1019
13	514.2	15.4	520	9	US-09-878-178-1019
14	514.2	15.4	520	9	US-10-146-502-1019
15	512.4	15.3	517	9	US-10-066-543-3247
16	511.4	15.3	627	9	US-09-871-161-514
17	511	15.3	512	9	US-10-066-543-2636
18	510.4	15.3	512	9	US-10-066-543-2512
19	508.8	15.2	512	9	US-10-060-036-2188

20	497	14.9	498	9	US-10-046-935-2092	Sequence 2092, Ap
21	497	14.9	498	9	US-09-878-178-2092	Sequence 2092, Ap
22	497	14.9	498	9	US-10-146-502-2092	Sequence 2092, Ap
23	490.4	14.4	496	9	US-10-060-036-2093	Sequence 2093, Ap
24	479.8	14.3	495	9	US-10-060-036-1950	Sequence 1950, Ap
25	469.2	14.0	492	9	US-10-046-935-571	Sequence 571, App
26	469.2	14.0	492	9	US-09-878-178-571	Sequence 571, App
27	469.2	14.0	492	9	US-10-146-502-571	Sequence 571, App
28	461.4	13.8	489	9	US-10-046-935-98	Sequence 98, App
29	461.4	13.8	489	9	US-09-878-178-98	Sequence 98, App
30	461.4	13.8	489	9	US-10-146-502-98	Sequence 98, App
31	450	13.5	450	10	US-09-998-598-696	Sequence 696, App
32	450	13.5	450	10	US-09-998-598-1002	Sequence 1002, App
33	445.2	13.3	450	10	US-09-998-598-1285	Sequence 1285, App
34	431.4	12.9	516	10	US-09-920-300A-409	Sequence 409, App
35	431.4	12.9	516	12	US-10-033-528-409	Sequence 409, App
36	425.6	12.7	451	9	US-10-066-543-1412	Sequence 1412, Ap
37	420.2	12.6	451	9	US-10-066-543-1413	Sequence 1413, Ap
38	410.4	12.3	568	10	US-09-920-300A-705	Sequence 705, App
39	410.4	12.3	568	12	US-10-033-528-705	Sequence 705, App
40	408.4	12.2	568	10	US-09-998-598-843	Sequence 843, App
41	401.4	12.0	429	9	US-10-066-543-1531	Sequence 1531, App
42	386.4	11.6	568	10	US-09-998-598-723	Sequence 723, App
43	367.4	11.0	375	9	US-10-066-543-2626	Sequence 2626, Ap
44	367	11.0	367	9	US-10-060-036-1960	Sequence 1960, Ap
45	367	11.0	367	9	US-10-060-036-2398	Sequence 2398, Ap

ALIGNMENTS

RESULT 1
US-10-025-380-1076
Sequence 1076, Application US/10025380
Publication No. US20020182191A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Steink, John A.
APPLICANT: Wang, Tongtong
APPLICANT: Jiang, Yuhui
APPLICANT: Smith, Carole L.
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Skeiky, Yasir A. W.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick Thomas S.
APPLICANT: Carter, Darrick
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.471C14
CURRENT APPLICATION NUMBER: US/10/025,380
CURRENT FILING DATE: 2001-12-19
NUMBER OF SEQ ID NOS: 1129
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1076
LENGTH: 3345
TYPE: DNA
ORGANISM: Homo sapiens
US-10-025-380-1076

Query Match 100.0%; Score 3345; DB 9; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GAATTCGCTTCGACCACTGAATGAAAGAAAGACTTTTAACACCAATTTGTGACTTA 60
DB 1 GAATTCGCTTCGACCACTGAATGAAAGAAAGACTTTTAACACCAATTTGTGACTTA 60

QY 61 CAGAAAGAAATTTGATTAAGAAAACTATGATCTTCAGGCCATCTTCACTCCCTGT 120
| | | | |
Db 61 CAGAAAGAAATTTGATTAAGAAAACTATGATCTTCAGGCCATCTTCACTCCCTGT 120
QY 121 CTTCTTAATGCTTTATTTGGCACTGATATGCGCAAGAGGGGAATTTAGTGAACCCCTG 180
| | | | |
Db 121 CTTCTTAATGCTTTATTTGGCACTGATATGCGCAAGAGGGGAATTTAGTGAACCCCTG 180
QY 181 AAACCCATGACATTTTCTATTTATGAAAGCCCAAGACCGAGTCAATTTATTTTCAATTT 240
| | | | |
Db 181 AAACCCATGACATTTTCTATTTATGAAAGCCCAAGACCGAGTCAATTTATTTTCAATTT 240
QY 241 AAGCCCAATCTCTGCTGATCTTTGAATTAATGAGGAGACAGACAACTATTTGTG 300
| | | | |
Db 241 AAGCCCAATCTCTGCTGATCTTTGAATTAATGAGGAGACAGACAACTATTTGTG 300
QY 301 ATAGAACGGAGGAGACTTGTATTAACAAGACCTTTGACAGAGAAACAAGATCTACT 360
| | | | |
Db 301 ATAGAACGGAGGAGACTTGTATTAACAAGACCTTTGACAGAGAAACAAGATCTACT 360
QY 361 CACATCTCCAGGTTTGCAGCCCTGAGCGCTAATGGAATTAATGAGAGGCTCCAGTCCCT 420
| | | | |
Db 361 CACATCTCCAGGTTTGCAGCCCTGAGCGCTAATGGAATTAATGAGAGGCTCCAGTCCCT 420
QY 421 ATCAACATAGAAAGTGAAGGACATCAACGACCAATGCAACCACTTTCTCCAGTCAAGTAC 480
| | | | |
Db 421 ATCAACATAGAAAGTGAAGGACATCAACGACCAATGCAACCACTTTCTCCAGTCAAGTAC 480
QY 481 GAAAGCTCAGTAAAGGAGAACTCTGCGCCAGAGAAAGCCCTTCTGTATGTCAATGCGACA 540
| | | | |
Db 481 GAAAGCTCAGTAAAGGAGAACTCTGCGCCAGAGAAAGCCCTTCTGTATGTCAATGCGACA 540
QY 541 GACCTGATGATCCGGGCACTCCCAATGCGCAAGCTTTATTAACGATTTGTATCCACTT 600
| | | | |
Db 541 GACCTGATGATCCGGGCACTCCCAATGCGCAAGCTTTATTAACGATTTGTATCCACTT 600
QY 541 GACCTGATGATCCGGGCACTCCCAATGCGCAAGCTTTATTAACGATTTGTATCCACTT 600
| | | | |
Db 541 GACCTGATGATCCGGGCACTCCCAATGCGCAAGCTTTATTAACGATTTGTATCCACTT 600
QY 601 CCCATGATCAACAAATGTATATTTCTTCAATCAACAAACGAGGACATCTCTT 660
| | | | |
Db 601 CCCATGATCAACAAATGTATATTTCTTCAATCAACAAACGAGGACATCTCTT 660
QY 661 ACCGAGAGGGATCTCAGGAATGAAATCTGTAAAGAACTCTTCTATTAATCTGTATC 720
| | | | |
Db 661 ACCGAGAGGGATCTCAGGAATGAAATCTGTAAAGAACTCTTCTATTAATCTGTATC 720
QY 721 TCAATGAAAGCAATGAGGAGGCGCAGAGTGAATTTCTTCAATGATTAACATCTGTGAT 780
| | | | |
Db 721 TCAATGAAAGCAATGAGGAGGCGCAGAGTGAATTTCTTCAATGATTAACATCTGTGAT 780
QY 781 ATCATAGTGAAGAGAAATTTTGGAAAGCAACAAACCTGTGAGATGTGGAATCTCA 840
| | | | |
Db 781 ATCATAGTGAAGAGAAATTTTGGAAAGCAACAAACCTGTGAGATGTGGAATCTCA 840
QY 841 ACTGATCTCAACCCCATCAAAATCACTCAGGTGCGGTGAATGATCCCGGTGCAATAT 900
| | | | |
Db 841 ACTGATCTCAACCCCATCAAAATCACTCAGGTGCGGTGAATGATCCCGGTGCAATAT 900
QY 901 TCTTAATGTTGCAAAAGAGAACTGCGCAAGATTTCCATTTTCAATTTGACAGAGAGAT 960
| | | | |
Db 901 TCTTAATGTTGCAAAAGAGAACTGCGCAAGATTTCCATTTTCAATTTGACAGAGAGAT 960
QY 961 ATTTACGTGACTCAGCCCTTGAACCGAGAGAAAGAGATGATATGTTTTTATGACAT 1020
| | | | |
Db 961 ATTTACGTGACTCAGCCCTTGAACCGAGAGAAAGAGATGATATGTTTTTATGACAT 1020
QY 1021 GCAAGAGTGAAGTACGAGAAACCACTTTCAATATCGCTGAAATTCATGTAAGATTA 1080
| | | | |
Db 1021 GCAAGAGTGAAGTACGAGAAACCACTTTCAATATCGCTGAAATTCATGTAAGATTA 1080
QY 1081 GATATTAATGATTAATCCACTTAATGTCTGCAACGATTAACCGATTTTGAAGTCAAGAG 1140
| | | | |
Db 1081 GATATTAATGATTAATCCACTTAATGTCTGCAACGATTAACCGATTTTGAAGTCAAGAG 1140
QY 1141 AATGAAGACTGGGTAAACGATATGCGGACCCCTTACTGACATGACAGGATGAAGAAAT 1200
| | | | |

Db 1141 AATGAAGACTGGGTAAACGATATGCGGACCCCTTACTGACATGACAGGATGAAGAAAT 1200
| | | | |
QY 1201 ACTGCCAAGATTTTCTAAATCTACAGATTTTGGACCAAACTCCCAATCCATGAT 1260
| | | | |
Db 1201 ACTGCCAAGATTTTCTAAATCTACAGATTTTGGACCAAACTCCCAATCCATGAT 1260
QY 1261 GGACTCTCTTAATCCAAACCTATGCTGGAATGTACAGTATGCTAAACAGTCTTGAAG 1320
| | | | |
Db 1261 GGACTCTCTTAATCCAAACCTATGCTGGAATGTACAGTATGCTAAACAGTCTTGAAG 1320
QY 1321 AAGCAAGATTAATCTCAGTACAACTTAACGATAGAGTGTCTGACAAAGATTTCAAGAC 1380
| | | | |
Db 1321 AAGCAAGATTAATCTCAGTACAACTTAACGATAGAGTGTCTGACAAAGATTTCAAGAC 1380
QY 1381 CTTTGTGTTGTGCAATTCACAGTATTAATGATTCAGATCCCACTCTTGAAGAA 1440
| | | | |
Db 1381 CTTTGTGTTGTGCAATTCACAGTATTAATGATTCAGATCCCACTCTTGAAGAA 1440
QY 1441 TCAGATTAATGAAACCTGACTCTTGTGGAAGACAAACATTTGGTCCACATCTTAAC 1500
| | | | |
Db 1441 TCAGATTAATGAAACCTGACTCTTGTGGAAGACAAACATTTGGTCCACATCTTAAC 1500
QY 1501 ATCCAGGCCACTGATCTGATGAGCCATTAATCTGGAAGTTCTAAATTCGTATCATATC 1560
| | | | |
Db 1501 ATCCAGGCCACTGATCTGATGAGCCATTAATCTGGAAGTTCTAAATTCGTATCATATC 1560
QY 1561 ATTAAGGAGACAGTGAAGGAGCGCTGGGGTGTGAACAGATCCCAATCCACACCGGA 1620
| | | | |
Db 1561 ATTAAGGAGACAGTGAAGGAGCGCTGGGGTGTGAACAGATCCCAATCCACACCGGA 1620
QY 1621 TATGTCAATTAATAAGCCCTTGAATTTGAAACAGAGCTGTTTCCAATGTGTTC 1680
| | | | |
Db 1621 TATGTCAATTAATAAGCCCTTGAATTTGAAACAGAGCTGTTTCCAATGTGTTC 1680
QY 1681 AAAGCAAGAAATCTGAGCCCTTGAATGTGTGTGATGAGTACAAATGATCTTTTGC 1740
| | | | |
Db 1681 AAAGCAAGAAATCTGAGCCCTTGAATGTGTGTGATGAGTACAAATGATCTTTTGC 1740
QY 1741 AAGTTCAAGCTTTATTTGACAGATGATGATGAAGACCTCAATTTTCCCAACGATATC 1800
| | | | |
Db 1741 AAGTTCAAGCTTTATTTGACAGATGATGATGAAGACCTCAATTTTCCCAACGATATC 1800
QY 1801 CAAGCGAAGTCAAGTGAAGTATGATGATGAGTAAAGTGGAGCAATGATCTGCGAAG 1860
| | | | |
Db 1801 CAAGCGAAGTCAAGTGAAGTATGATGATGAGTAAAGTGGAGCAATGATCTGCGAAG 1860
QY 1861 GATCCAGAAAGTCTGACATTAAGCTATTCATCTGAGGAGGAGACAAAGAGTTGGCTTAA 1920
| | | | |
Db 1861 GATCCAGAAAGTCTGACATTAAGCTATTCATCTGAGGAGGAGACAAAGAGTTGGCTTAA 1920
QY 1921 ATTGACCAAGTATCTGTGATGATCTTTAATGTGTGCTCAATTTGACAGAGAGCGGAAT 1980
| | | | |
Db 1921 ATTGACCAAGTATCTGTGATGATCTTTAATGTGTGCTCAATTTGACAGAGAGCGGAAT 1980
QY 1981 CCAATATCGGTTAACAAGTGTGAGGCGCAGAGAAAGTGAAGGAGGCTCTTAAAGCTGTGTCA 2040
| | | | |
Db 1981 CCAATATCGGTTAACAAGTGTGAGGCGCAGAGAAAGTGAAGGAGGCTCTTAAAGCTGTGTCA 2040
QY 2041 GAGTTCCACTGATCTCTTATGATGATGATGATCAACCTTCCAGGCTAAGCAAGATAC 2100
| | | | |
Db 2041 GAGTTCCACTGATCTCTTATGATGATGATGATCAACCTTCCAGGCTAAGCAAGATAC 2100
QY 2101 ACGGGCTTGTCTTCTGCAATCCCTCAAGTGAACCTGGAAGTCTAATTTTCAGGCTACT 2160
| | | | |
Db 2101 ACGGGCTTGTCTTCTGCAATCCCTCAAGTGAACCTGGAAGTCTAATTTTCAGGCTACT 2160
QY 2161 GATGATGATGAGCACTTAATTTGGGGGCTCCCAATTTTCAATTTTCCCTCGGCAAGTGAAG 2220
| | | | |
Db 2161 GATGATGATGAGCACTTAATTTGGGGGCTCCCAATTTTCAATTTTCCCTCGGCAAGTGAAG 2220
QY 2221 TTACAAAACGACTGGGAGTTTCCAAATCAATGATCTCATGCTCCGACTGTCAACAG 2280
| | | | |

Db 2221 TTACAAAACGACTGGAAAGTTTCCAAATATCATGTACTATGCCCACTGTACACAG 2280
Qy 2281 CACACAGACTTGGAGAGAGGCGTATGCTCTTGAATCCGATCAATGATGGGGTCCG 2340
Db 2281 CACACAGACTTGGAGAGAGGCGTATGCTCTTGAATCCGATCAATGATGGGGTCCG 2340
Qy 2341 CCACCCCTTGGAGAGGCAATGTTCTTTTACAGTTACATTTCTGCACTGTGTGGAAGAAGT 2400
Db 2341 CCACCCCTTGGAGAGGCAATGTTCTTTTACAGTTACATTTCTGCACTGTGTGGAAGAAGT 2400
Qy 2401 TGTTCGGGCGAGAGGTCACACAGCTGGGATACCCCTGCGGACATGGCACTGTGTGTA 2460
Db 2401 TGTTCGGGCGAGAGGTCACACAGCTGGGATACCCCTGCGGACATGGCACTGTGTGTA 2460
Qy 2461 CTGCTGACCAACCTTCTGTGATGTATTAATTTTACAGTTGTGTATCCGATTAAG 2520
Db 2461 CTGCTGACCAACCTTCTGTGATGTATTAATTTTACAGTTGTGTATCCGATTAAG 2520
Qy 2521 AAGGATTAAGGCAAGATATGTTGAAAGTCTCAGCATCTGAAGTCAAACTCTGAGA 2580
Db 2521 AAGGATTAAGGCAAGATATGTTGAAAGTCTCAGCATCTGAAGTCAAACTCTGAGA 2580
Qy 2581 AGCGATTTGAAAGAAAGATTTGTAATTTATATAGCAAGTGTATTTAGCAACAACA 2640
Db 2581 AGCGATTTGAAAGAAAGATTTGTAATTTATATAGCAAGTGTATTTAGCAACAACA 2640
Qy 2641 TCTCATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCT 2700
Db 2641 TCTCATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCT 2700
Qy 2701 TGTCTTTTAAATTTTGTAAATATTTCTTTTGTAGGTGAGTCTCTGTGCCAG 2760
Db 2701 TGTCTTTTAAATTTTGTAAATATTTCTTTTGTAGGTGAGTCTCTGTGCCAG 2760
Qy 2761 GCTGAGATAGAGGTGTATCCAGCTCACTGACACTCCGCTCTGTGGTTTACATGA 2820
Db 2761 GCTGAGATAGAGGTGTATCCAGCTCACTGACACTCCGCTCTGTGGTTTACATGA 2820
Qy 2821 TTTCTCTGCTCAGCTTCTTATAGAGTGTGTATAGGACCCACACATGCTCCAGCT 2880
Db 2821 TTTCTCTGCTCAGCTTCTTATAGAGTGTGTATAGGACCCACACATGCTCCAGCT 2880
Qy 2881 AATTTTGTATTTTAAATAGAGCGGGTTTCCGCAATTTGGCCAGGCTGTGTGAAC 2940
Db 2881 AATTTTGTATTTTAAATAGAGCGGGTTTCCGCAATTTGGCCAGGCTGTGTGAAC 2940
Qy 2941 CTGACGTCAGATGCTGCTGCTGTGTGTCTCCCAATACAGGATGAACCACTGACCCA 3000
Db 2941 CTGACGTCAGATGCTGCTGCTGTGTGTCTCCCAATACAGGATGAACCACTGACCCA 3000
Qy 3001 CCTACTAGATTTTCTGCTATAGACATTTAGAGATTTTCAATTTTCCATGACAT 3060
Db 3001 CCTACTAGATTTTCTGCTATAGACATTTAGAGATTTTCAATTTTCCATGACAT 3060
Qy 3061 TTTTCTCTGCAATAGGCTTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120
Db 3061 TTTTCTCTGCAATAGGCTTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120
Qy 3121 CATTAATATTTCTGATATTTTCTTATACAGAGATATATCATGTTGTCTCATAGA 3180
Db 3121 CATTAATATTTCTGATATTTTCTTATACAGAGATATATCATGTTGTCTCATAGA 3180
Qy 3181 ACTGCTGATTCATTTATGTTTCTGATTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGT 3240
Db 3181 ACTGCTGATTCATTTATGTTTCTGATTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGT 3240
Qy 3241 TCTTTGTATTTCTACTGAATTTCAACATTTGTGAGAGAGAGAGAGAGAGAGAGAGAG 3300
Db 3241 TCTTTGTATTTCTACTGAATTTCAACATTTGTGAGAGAGAGAGAGAGAGAGAGAGAG 3300
Qy 3301 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3345
Db 3301 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3345

RESULT 2
US-09-922-217-1076
Sequence 1076, Application US/09922217
Patent No. US2002007641A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Stolk, John A.
APPLICANT: Wang, Tonglong
APPLICANT: Jiang, Yugu
APPLICANT: Smith, Carole Lynn
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.471C13
CURRENT APPLICATION NUMBER: US/09/922,217
NUMBER OF FILING DATE: 2001-08-03
NUMBER OF SEQ ID NOS: 1124
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1076
LENGTH: 3345
TYPE: DNA
ORGANISM: Homo sapiens
US-09-922-217-1076

Query Match 100.0%; Score 3345; DB 10; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAAATTCGCTCGACCACTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db 1 GAAATTCGCTCGACCACTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Qy 61 CAGAAAGAAATTTGAAATTAAGAAACTATGATATCTTACAGGCCATCTTCCCTGTGT 120
Db 61 CAGAAAGAAATTTGAAATTAAGAAACTATGATATCTTACAGGCCATCTTCCCTGTGT 120
Qy 121 CTCTTATGCTTATTTTGGCACTGATATGCTGACAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 121 CTCTTATGCTTATTTTGGCACTGATATGCTGACAGAGAGAGAGAGAGAGAGAGAGAG 180
Qy 181 AAACCATGACATTTTCTATTTATGAGGCAAGACGAGTCAATTAATTTCCAGTTT 240
Db 181 AAACCATGACATTTTCTATTTATGAGGCAAGACGAGTCAATTAATTTCCAGTTT 240
Qy 241 AAGGCCAATCTCTCTGCTGTGATTTTGAATCTGAGGAGAGAGAGAGAGAGAGAGAG 300
Db 241 AAGGCCAATCTCTCTGCTGTGATTTTGAATCTGAGGAGAGAGAGAGAGAGAGAGAG 300
Qy 301 ATGAGACGAGAGAGAGATTTCTGATTTACACAGGCTTTGAGAGAGAGAGAGAGAGAG 360
Db 301 ATGAGACGAGAGAGAGATTTCTGATTTACACAGGCTTTGAGAGAGAGAGAGAGAGAG 360
Qy 361 CACATCTCAGGTTGACAGCCCTGAGAGCTTAATGAAATTAATGAGAGAGAGAGAGAG 420
Db 361 CACATCTCAGGTTGACAGCCCTGAGAGCTTAATGAAATTAATGAGAGAGAGAGAGAG 420
Qy 421 ATCACCATGAG 480
Db 421 ATCACCATGAG 480
Qy 481 GAAAGCTCAGTAAAG 540
Db 481 GAAAGCTCAGTAAAG 540
Qy 541 GACCTGATGATCCGCGCACTCCCAATGAGCAGCTTATATACAGATGTATCCAGCTT 600

Db 541 GACCTGATGATCGCGCACTCCCAATGGCCAGCTTTATTAACGATTTGTCATCACTT 600
Qy 601 CCCATGATCAACAATGTCATGTAATCTTTCAGATCAACAACAAACGGAGCATCTCTT 660
Db 601 CCCATGATCAACAATGTCATGTAATCTTTCAGATCAACAACAAACGGAGCATCTCTT 660
Qy 661 ACCCGAAGGGATCTCAGGAATTTGAATCCGTGAAGATCCCTTAATCTGGATC 720
Db 661 ACCCGAAGGGATCTCAGGAATTTGAATCCGTGAAGATCCCTTAATCTGGATC 720
Qy 721 TCAATGAGGACATGAGGAGCCAGAGTGAATTCCTTCAGTGAATACCAATCTGTGAT 780
Db 721 TCAATGAGGACATGAGGAGCCAGAGTGAATTCCTTCAGTGAATACCAATCTGTGAT 780
Qy 781 ATCATAGTACAGAAATATTTGGAAAGCACAAACCTGTGAGATGTGGAAACTCA 840
Db 781 ATCATAGTACAGAAATATTTGGAAAGCACAAACCTGTGAGATGTGGAAACTCA 840
Qy 841 ACTGATCTCAACCCCATCAAAATCACTCAGGTGGGTGGAAATGATCCGGTGACATAT 900
Db 841 ACTGATCTCAACCCCATCAAAATCACTCAGGTGGGTGGAAATGATCCGGTGACATAT 900
Qy 901 TCCTTAGTTGCAAGAAGAGCTGCCAAGATTTCCATTTTCATTTGACAGAGAGAT 960
Db 901 TCCTTAGTTGCAAGAAGAGCTGCCAAGATTTCCATTTTCATTTGACAGAGAGAT 960
Qy 961 ATTTAGTACTCAGCCCTTGAGCCGAGAGAAAGATGCAATGTTTTTATGCACTT 1020
Db 961 ATTTAGTACTCAGCCCTTGAGCCGAGAGAAAGATGCAATGTTTTTATGCACTT 1020
Qy 1021 GCAAGAGATGATGCGGAAACCACTTCAATCCGCTGGAATTCATGTAAAGTTAA 1080
Db 1021 GCAAGAGATGATGCGGAAACCACTTCAATCCGCTGGAATTCATGTAAAGTTAA 1080
Qy 1081 GATATTAATGATTAATCCACTACATGTCCTGACACAGTAACCGTAATTTGAGTCA 1140
Db 1081 GATATTAATGATTAATCCACTACATGTCCTGACACAGTAACCGTAATTTGAGTCA 1140
Qy 1141 AATTAAGATCTGGGTAACATTAAGGAGCCCTTACTGCAATGACAGGAGTAAAGAAAT 1200
Db 1141 AATTAAGATCTGGGTAACATTAAGGAGCCCTTACTGCAATGACAGGAGTAAAGAAAT 1200
Qy 1201 ACTGCAAGATTTCTTAACCTTAAGATGTGGAGCAATCCCAATCTTCCATGAT 1260
Db 1201 ACTGCAAGATTTCTTAACCTTAAGATGTGGAGCAATCCCAATCTTCCATGAT 1260
Qy 1261 GGAATCTTCTTAATCCAACTATGCTGGAATGTTACGTTAAGCTTAACGCTTGAAG 1320
Db 1261 GGAATCTTCTTAATCCAACTATGCTGGAATGTTACGTTAAGCTTAACGCTTGAAG 1320
Qy 1321 AAGCAAGATCTCTCAGTACATTTAACATTAAGATGTGTGACAAAGATTTCAAGACC 1380
Db 1321 AAGCAAGATCTCTCAGTACATTTAACATTAAGATGTGTGACAAAGATTTCAAGACC 1380
Qy 1381 CTTGTTTGTGCAATCAAGCTTAATGATTAATGATCAATGATCCCATCTTTGAATA 1440
Db 1381 CTTGTTTGTGCAATCAAGCTTAATGATTAATGATCAATGATCCCATCTTTGAATA 1440
Qy 1441 TCAGATTAATGAAACCTGACTCTTGCTGAGACAAACATTTGGGTCCACATCTTAAC 1500
Db 1441 TCAGATTAATGAAACCTGACTCTTGCTGAGACAAACATTTGGGTCCACATCTTAAC 1500
Qy 1501 ATCAAGGCACTGATGCTGATGAGCCATTTACTGGAGTTCTAAATTTCTGTATCATTC 1560
Db 1501 ATCAAGGCACTGATGCTGATGAGCCATTTACTGGAGTTCTAAATTTCTGTATCATTC 1560
Qy 1561 ATTAAGGAGACAGTGGGAGGCGCTGGGGGTGACACAGATCCCATACCAACCGGA 1620
Db 1561 ATTAAGGAGACAGTGGGAGGCGCTGGGGGTGACACAGATCCCATACCAACCGGA 1620
Qy 1621 TATGTCATTAATTAAGGCTCTTGATTTTGAACAGAGCTGTTTCCATTTGTGTC 1680
Db 1621 TATGTCATTAATTAAGGCTCTTGATTTTGAACAGAGCTGTTTCCATTTGTGTC 1680
Db 1621 TATGTCATTAATTAAGGCTCTTGATTTTGAACAGAGCTGTTTCCATTTGTGTC 1680
Qy 1681 AAGCAGAAATCTGAGCCTAGTGTGTTGGTGAAGTACAAAGTATCTTTTGGC 1740
Db 1681 AAGCAGAAATCTGAGCCTAGTGTGTTGGTGAAGTACAAAGTATCTTTTGGC 1740
Qy 1741 AAGTCAAGCTTAATGTCAGATGTAATGAAGACCTCAATTTTCCAAACGATTC 1800
Db 1741 AAGTCAAGCTTAATGTCAGATGTAATGAAGACCTCAATTTTCCAAACGATTC 1800
Qy 1801 CAGCCAAAGTCAATGAGATGTAATGAGTACATTAAGGATGAGTATGATGATCCCAAG 1860
Db 1801 CAGCCAAAGTCAATGAGATGTAATGAGTACATTAAGGATGAGTATGATGATCCCAAG 1860
Qy 1861 GATCCAGAAAGTGTGACATTAAGCTTAATCACTGAGGAGAGACAAAGAGTGGCTTAA 1920
Db 1861 GATCCAGAAAGTGTGACATTAAGCTTAATCACTGAGGAGAGACAAAGAGTGGCTTAA 1920
Qy 1921 ATTGACCACTGACTGATGATCTTTAGTGTGCTCAATTTGACAGAGAGCCGAAAT 1980
Db 1921 ATTGACCACTGACTGATGATCTTTAGTGTGCTCAATTTGACAGAGAGCCGAAAT 1980
Qy 1981 CCAATGCGGTACAGTGTGCTGACAGAAATGAGGAGGCTTCTTAAGCTGTGTCA 2040
Db 1981 CCAATGCGGTACAGTGTGCTGACAGAAATGAGGAGGCTTCTTAAGCTGTGTCA 2040
Qy 2041 GAGTTCACCTGATCTCTTAATGATGTAATGACAAACCTCCAGGCTAGCCAGACTAC 2100
Db 2041 GAGTTCACCTGATCTCTTAATGATGTAATGACAAACCTCCAGGCTAGCCAGACTAC 2100
Qy 2101 ACGGGCTTGTCTTCTGCAATCCCTCAGTGCACCTGGAAGTCTATTTTGAAGCTACT 2160
Db 2101 ACGGGCTTGTCTTCTGCAATCCCTCAGTGCACCTGGAAGTCTATTTTGAAGCTACT 2160
Qy 2161 GATGATGATAGACATTAATTTGGGGTCCCATTTTACATTTTCCCTGGCAGTGAAGC 2220
Db 2161 GATGATGATAGACATTAATTTGGGGTCCCATTTTACATTTTCCCTGGCAGTGAAGC 2220
Qy 2221 TTAACAAAGACTGGGAGGTTTCCAAATCAATGATGATGATGATGATGATGATGATGAT 2280
Db 2221 TTAACAAAGACTGGGAGGTTTCCAAATCAATGATGATGATGATGATGATGATGATGAT 2280
Qy 2281 CACACAGACTTTGAGGAGGAGGCTGATGTCCTTGAATCCGATGATGATGATGATGAT 2340
Db 2281 CACACAGACTTTGAGGAGGAGGCTGATGTCCTTGAATCCGATGATGATGATGATGAT 2340
Qy 2341 CCACTCTGGAAGGATGTTTCTTAACAGTTAATCTGCAATGATGATGATGATGATGAT 2400
Db 2341 CCACTCTGGAAGGATGTTTCTTAACAGTTAATCTGCAATGATGATGATGATGATGAT 2400
Qy 2401 TGTTCGCGCAGCAGGTCACAGATGAGGATCCCACTGAGGATGAGGATGAGGATGAG 2460
Db 2401 TGTTCGCGCAGCAGGTCACAGATGAGGATCCCACTGAGGATGAGGATGAGGATGAG 2460
Qy 2461 CTGCTGACCACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
Db 2461 CTGCTGACCACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
Qy 2521 AAGATTAAGGCAAGATTAATGTTAAAGTCTCAAGCATCTGAATCTGAGAC 2580
Db 2521 AAGATTAAGGCAAGATTAATGTTAAAGTCTCAAGCATCTGAATCTGAGAC 2580
Qy 2581 AGCTGAATTTGAAGAAAGTGTGAATTTAATGAGTGAATGAGTGAATGAGTGAATGAG 2640
Db 2581 AGCTGAATTTGAAGAAAGTGTGAATTTAATGAGTGAATGAGTGAATGAGTGAATGAG 2640
Qy 2641 TCTCATCTTAATCTTTTCACTAAGCTGATTAATTTTAAACAGATATTCCTCT 2700
Db 2641 TCTCATCTTAATCTTTTCACTAAGCTGATTAATTTTAAACAGATATTCCTCT 2700
Qy 2701 TGTCTTTAATTAATTTGCTTAATTTCTTTTGAAGTGAAGTCTGTGCGCCAG 2760
Db 2701 TGTCTTTAATTAATTTGCTTAATTTCTTTTGAAGTGAAGTCTGTGCGCCAG 2760

QY 2761 GCTGAGTACAGTGTGTGATCCAGCTCACTGCACTCCGCTCCGCTGAGTTTCACATGA 2820
Db 2761 GCTGAGTACAGTGTGTGATCCAGCTCACTGCACTCCGCTCCGCTGAGTTTCACATGA 2820
QY 2821 TTCTCCCTGCTCAGCTTCTTAAGTACTGGTTTACAGGCAACCAACCACTGCTGCT 2880
Db 2821 TTCTCCCTGCTCAGCTTCTTAAGTACTGGTTTACAGGCAACCAACCACTGCTGCT 2880
QY 2881 AATTTTGTATTTTAAATAGAGAGAGGTTTGGCATTTGGCCAGGCTGGCTTGAATC 2940
Db 2881 AATTTTGTATTTTAAATAGAGAGAGGTTTGGCATTTGGCCAGGCTGGCTTGAATC 2940
QY 2941 CTGACGTCAAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
Db 2941 CTGACGTCAAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
QY 3001 CCTACTAGATATTTTCAATGTGCTATAGACATTTAGAGATTTTTCATTTTCCATGACAT 3060
Db 3001 CCTACTAGATATTTTCAATGTGCTATAGACATTTAGAGATTTTTCATTTTCCATGACAT 3060
QY 3061 TTTTCTCTGCAAAATGGCTAGCTACTGTTTTCCTTTTGGGGCAAGACACT 3120
Db 3061 TTTTCTCTGCAAAATGGCTAGCTACTGTTTTCCTTTTGGGGCAAGACACT 3120
QY 3121 CATTAATATTTCTGATATTTTCTTATCAAGAGATATATCATGTTGTCTCATAGA 3180
Db 3121 CATTAATATTTCTGATATTTTCTTATCAAGAGATATATCATGTTGTCTCATAGA 3180
QY 3181 ACTGCTGATTCATTTATGTTTTCATTTTCATCTGCTGCTGCTGCTGCTGCTGCTGAC 3240
Db 3181 ACTGCTGATTCATTTATGTTTTCATTTTCATCTGCTGCTGCTGCTGCTGCTGCTGAC 3240
QY 3241 TCCCTTGTATTTCTGATATTTTCAACATTTTGTCAAGAGAGAGAGAGAGAGAGAGAG 3300
Db 3241 TCCCTTGTATTTCTGATATTTTCAACATTTTGTCAAGAGAGAGAGAGAGAGAGAGAG 3300
QY 3301 GAAAAATTAATTAATAAAG 3345
Db 3301 GAAAAATTAATTAATAAAG 3345

RESULT 3
US-09-962-436-302
; Sequence 302, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppel, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 302
; LENGTH: 3345
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-302

Query Match 100.0%; Score 3345; DB 10; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGCTGCTGACCACTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db 1 GAATTCGCTGCTGACCACTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60

QY 61 CAGAAAGAAATTTGAATTAAGAAAGAACTATGATCTTCAAGGCCATCTTCACTCCCTGTGT 120
Db 61 CAGAAAGAAATTTGAATTAAGAAAGAACTATGATCTTCAAGGCCATCTTCACTCCCTGTGT 120
QY 121 CTCTTATGCTTTTATTTGGCACTGATATGCGCAAGAGGGGAACTTTAGTGAACCTGT 180
Db 121 CTCTTATGCTTTTATTTGGCACTGATATGCGCAAGAGGGGAACTTTAGTGAACCTGT 180
QY 181 AAACCCATGACATTTTCTATTTATGAAGGCAAGAACCGGATCAATTAATTAATTCAGATT 240
Db 181 AAACCCATGACATTTTCTATTTATGAAGGCAAGAACCGGATCAATTAATTAATTCAGATT 240
QY 241 AAGGCCAATCTCTCTGTGATCTTTGAACTAATCTGGGAGACAGACAAATATTTGTG 300
Db 241 AAGGCCAATCTCTCTGTGATCTTTGAACTAATCTGGGAGACAGACAAATATTTGTG 300
QY 301 ATGGAACGGAG 360
Db 301 ATGGAACGGAG 360
QY 361 CACAATCTCCAGGTTGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 361 CACAATCTCCAGGTTGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 ATCACAATGAG 480
Db 421 ATCACAATGAG 480
QY 481 GAGGCTCAGTAAAG 540
Db 481 GAGGCTCAGTAAAG 540
QY 541 GACCTGAGATATCCGAG 600
Db 541 GACCTGAGATATCCGAG 600
QY 601 CCCATGATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db 601 CCCATGATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 661 ACCGAGAGAGAGATCTGAG 720
Db 661 ACCGAGAGAGAGATCTGAG 720
QY 721 TCAGTAAAG 780
Db 721 TCAGTAAAG 780
QY 781 ATCATAGTGAAGAGAGATATTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 ATCATAGTGAAGAGAGATATTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 ACTGATCTTCAACCCATCAAAATCACTCAGGTGCGGTGAATGATCCCGGTGACAAATAT 900
Db 841 ACTGATCTTCAACCCATCAAAATCACTCAGGTGCGGTGAATGATCCCGGTGACAAATAT 900
QY 901 TCCCTTGTGAG 960
Db 901 TCCCTTGTGAG 960
QY 961 ATTTAGTGAAG 1020
Db 961 ATTTAGTGAAG 1020
QY 1021 GCAAG 1080
Db 1021 GCAAG 1080
QY 1081 GATATTAATGATATATCACTGATATGATGATGATGATGATGATGATGATGATGATGATG 1140
Db 1081 GATATTAATGATATATCACTGATATGATGATGATGATGATGATGATGATGATGATGATG 1140
QY 1141 AATGAAG 1200

Db	1141	AATGAAAGCACTGGGTAACGTAATGGGAGCCCTTACTGCACATGACAGGAAATGAAAGAAAT	1200
Qy	1201	ACTGCCAACAGATTTTCTAAACTTAACAGATTTGTGAGCAAACTCCCAACTTCCATGGAT	1260
Db	1201	ACTGCMAACAGTTTCTTCTAACTTAACAGATTTGTGAGCAAACTCCCAACTTCCATGGAT	1260
Qy	1261	GGACTCTTCCATATCCAAACCTATGCTGGAAATGTTACAGTTAAGCTAAACAGTCCCTTGAAG	1320
Db	1261	GGACTCTTCCATATCCAAACCTATGCTGGAAATGTTACAGTTAAGCTAAACAGTCCCTTGAAG	1320
Qy	1321	AAGCAAGATACCTCCCTCAGTACAACCTTAAACGATAGGAGTGTGACAAAGATTTCAAGACC	1380
Db	1321	AAGCAAGATACCTCCCTCAGTACAACCTTAAACGATAGGAGTGTGACAAAGATTTCAAGACC	1380
Qy	1381	CTTTGTGTTGTGCAAAATCAACGTTAATGATATCAATGATCAGATCCCATCTTTGAAAAA	1440
Db	1381	CTTTGTGTTGTGCAAAATCAACGTTAATGATATCAATGATCAGATCCCATCTTTGAAAAA	1440
Qy	1441	TCAGATTATGGAACCTGACCTCTTGCTGTAAGACAAACATTTGGGTCCACATCTTAAAC	1500
Db	1441	TCAGATTATGGAACCTGACCTCTTGCTGTAAGACAAACATTTGGGTCCACATCTTAAAC	1500
Qy	1501	ATCCAGGCCACATGATGCTGATGAGCCATTACTGGGAGTTCTTAAATTCGTATCATATTC	1560
Db	1501	ATCCAGGCCACATGATGCTGATGAGCCATTACTGGGAGTTCTTAAATTCGTATCATATTC	1560
Qy	1561	ATTAAGGGAGACAGTGAAGGAGCGCCTGGGGGTTGACACAGATCCCATACCAACCGGA	1620
Db	1561	ATTAAGGGAGACAGTGAAGGAGCGCCTGGGGGTTGACACAGATCCCATACCAACCGGA	1620
Qy	1621	TATGTCATTAATTAAGAACCTCTTGATTTTGAACAGACGCTGTTTCCAACTGTGTTC	1680
Db	1621	TATGTCATTAATTAAGAACCTCTTGATTTTGAACAGACGCTGTTTCCAACTGTGTTC	1680
Qy	1681	AAAGCAGAAATCCCGAGCCTCTGATGTTGTGTGTGAAGTACAAATGCAAGTCTTTTGGC	1740
Db	1681	AAAGCAGAAATCCCGAGCCTCTGATGTTGTGTGTGAAGTACAAATGCAAGTCTTTTGGC	1740
Qy	1741	AAGTTCAACGTTAATTTGTGACAGATGTGATGAAGACCTCAATTTTCCCAACAGTATTC	1800
Db	1741	AAGTTCAACGTTAATTTGTGACAGATGTGATGAAGACCTCAATTTTCCCAACAGTATTC	1800
Qy	1801	CAGCGAAAGTCAGTGAAGATGTAGCTTAAGGCACTAAAGTGGCCAAATGTACATGCGAAG	1860
Db	1801	CAGCGAAAGTCAGTGAAGATGTAGCTTAAGGCACTAAAGTGGCCAAATGTACATGCGAAG	1860
Qy	1861	GATCCAGAAAGTCTGGACATTAAGCATTTCACTGAGGGGAGACACAAAGATGTGGCTTAA	1920
Db	1861	GATCCAGAAAGTCTGGACATTAAGCATTTCACTGAGGGGAGACACAAAGATGTGGCTTAA	1920
Qy	1921	ATTGACACAGTGACTGTGAGATCTTTAAGTGTGCTCCATTGAGACAGAGAAACCGGAAGT	1980
Db	1921	ATTGACACAGTGACTGTGAGATCTTTAAGTGTGCTCCATTGAGACAGAGAAACCGGAAGT	1980
Qy	1981	CCATATCGGGGTACAAAGTGTGGCCACAAAGTAAAGGGGGGTCTTCTTAAAGCTGTGTCA	2040
Db	1981	CCATATCGGGGTACAAAGTGTGGCCACAAAGTAAAGGGGGGTCTTCTTAAAGCTGTGTCA	2040
Qy	2041	GAGTTCCACCTGATCTTATGATGTGAATGACAACTCCACAGGCTAAGCCAAAGACATAC	2100
Db	2041	GAGTTCCACCTGATCTTATGATGTGAATGACAACTCCACAGGCTAAGCCAAAGACATAC	2100
Qy	2101	ACGGGCTTGTCTTCTGACATCCCTCAGTGCACCTGGAAGTCTCAATTTTGAAGGCTACT	2160
Db	2101	ACGGGCTTGTCTTCTGACATCCCTCAGTGCACCTGGAAGTCTCAATTTTGAAGGCTACT	2160
Qy	2161	GATGATGATCAGACTTAATTTTGGGGGTCCCAATTTTACATTTTCCCTGCGAGTGAAGC	2220
Db	2161	GATGATGATCAGACTTAATTTTGGGGGTCCCAATTTTACATTTTCCCTGCGAGTGAAGC	2220
Qy	2221	TTACAAAACGACTGGGAAGTTCCAAATTCATGTACTCATGCCCACTGTCTACAGG	2280

Db	2221	TTACAAAACACCTGGGAAGTTTCCAAAATCAATGATCTCATGCGCCACAGTCTACACAGG	2280
Qy	2281	CACACACACTTTGAGGAGAGAGGGGTATAGTGGTCTGTGATCCGGACATCAATGATGGGGGTCCGG	2340
Db	2281	CACACACACTTTGAGGAGAGAGGGGTATAGTGGTCTGTGATCCGGACATCAATGATGGGGGTCCGG	2340
Qy	2341	CCACCCTTGGAAAGGCATTTGTTTCTTTTCCAGTTACATTTCTGCAGTTGTGTGGAAAGAAGT	2400
Db	2341	CCACCCTTGGAAAGGCATTTGTTTCTTTTCCAGTTACATTTCTGCAGTTGTGTGTGGAAAGAAGT	2400
Qy	2401	TGTTTCCGGCCAGAGGTCAACAGACTGGGATACCACTGTGGGCATGGCAATTGGTATA	2460
Db	2401	TGTTTCCGGCCAGAGGTCAACAGACTGGGATACCACTGTGGGCATGGCAATTGGTATA	2460
Qy	2461	CTGTGTACCAACCTCTTGGTGATTTGGATATATTTTACAGTTGGTTATCCGATATAAG	2520
Db	2461	CTGTGTACCAACCTCTTGGTGATTTGGATATATTTTACAGTTGGTTATCCGATATAAG	2520
Qy	2521	AAGGATTAAGGCAMAAGATTAATGTTGAAGGCTCAACATCTGAAAGTCAAAACCTCTGAGA	2580
Db	2521	AAGGATTAAGGCAMAAGATTAATGTTGAAGGCTCAACATCTGAAAGTCAAAACCTCTGAGA	2580
Qy	2581	AGCTGAATTTGAAAAAGGAATGTTGAAATTTATATAGCAAGTCTATTTCAGCAACAACA	2640
Db	2581	AGCTGAATTTGAAAAAGGAATGTTGAAATTTATATAGCAAGTCTATTTCAGCAACAACA	2640
Qy	2641	TCTCATCTCTATTTACTTTTCAATCTTACAGTGCATTAATTTTTTAAACAGATATTCCTCT	2700
Db	2641	TCTCATCTCTATTTACTTTTCAATCTTACAGTGCATTAATTTTTTAAACAGATATTCCTCT	2700
Qy	2701	TGTCCTTTAATATTTTGTGTAATAATTTCTTTTGTGAGTGAGTCTTGTCTCTGTGCCAG	2760
Db	2701	TGTCCTTTAATATTTTGTGTAATAATTTCTTTTGTGAGTGAGTCTTGTCTCTGTGCCAG	2760
Qy	2761	GCTGAGATACAGTGGTGTGATCCAGCTCACTGCMACTCCGCTCTGTGGTTTCAATGA	2820
Db	2761	GCTGAGATACAGTGGTGTGATCCAGCTCACTGCMACTCCGCTCTGTGGTTTCAATGA	2820
Qy	2821	TTCTCTGCTCAGCTTCTTAAGTACTGGGTTTACAGGCCAACCAACCACTGCCAGCT	2880
Db	2821	TTCTCTGCTCAGCTTCTTAAGTACTGGGTTTACAGGCCAACCAACCACTGCCAGCT	2880
Qy	2881	AATTTTGTATTTTAAATAGAGAGGGGTTTGGCATTTGGCCAGGGCTGGTCTTGAATC	2940
Db	2881	AATTTTGTATTTTAAATAGAGAGGGGTTTGGCATTTGGCCAGGGCTGGTCTTGAATC	2940
Qy	2941	CTGACGTCAGATGATCTGCTGCTTGGTCTTCCCAATACAGGCATGAACACTGCACCCA	3000
Db	2941	CTGACGTCAGATGATCTGCTGCTTGGTCTTCCCAATACAGGCATGAACACTGCACCCA	3000
Qy	3001	CTTACTTATGATTTTCAATGTGTCTATPACATTAAGAGATTTTTCATTTTCCATGACAT	3060
Db	3001	CTTACTTATGATTTTCAATGTGTCTATPACATTAAGAGATTTTTCATTTTCCATGACAT	3060
Qy	3061	TTTTTCCTCTCGCAAAAGGCTTACTACTGTGTGTTTTTCCCTTTTGGGGCAAGACAGACT	3120
Db	3061	TTTTTCCTCTCGCAAAAGGCTTACTACTGTGTGTTTTTCCCTTTTGGGGCAAGACAGACT	3120
Qy	3121	CATTAAATATTTCTGTACATTTTTTCTTTATCAAGAGATATATCATGATGTGTCTCATAGA	3180
Db	3121	CATTAAATATTTCTGTACATTTTTTCTTTATCAAGAGATATATCATGATGTGTCTCATAGA	3180
Qy	3181	ACTGCGTGATTCATTTAATGTTTTTCTGATTCATCTGTGTCTTCTCATCTCTTGAC	3240
Db	3181	ACTGCGTGATTCATTTAATGTTTTTCTGATTCATCTGTGTCTTCTCATCTCTTGAC	3240
Qy	3241	TCTTTTGGTATTTTCACTGCAATTTTCAACATTTTGTCAAGAGAAAGAAAGTGAAGACTCAG	3300
Db	3241	TCTTTTGGTATTTTCACTGCAATTTTCAACATTTTGTCAAGAGAAAGAAAGTGAAGACTCAG	3300
Qy	3301	GAAGAAATTAATTAATTAAGAAAGAGCCTTTGGGGCCGCGAATTC	3345
Db	3301	GAAGAAATTAATTAATTAAGAAAGAGCCTTTGGGGCCGCGAATTC	3345

```

RESULT 4
US-09-833-263-1076
; Sequence 1076, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; NUMBER OF SEQ ID NOS: 2001-04-10
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 1076
; LENGTH: 3345
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-263-1076

Query Match      100.0%; Score 3345; DB 10; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGCTCCGACCTGATGGAAGAAAGACCTTTTAAACCACCATTTTGTGACTTA 60
DB 1 GAATTCGCTCCGACCTGATGGAAGAAAGACCTTTTAAACCACCATTTTGTGACTTA 60
QY 61 CAGAAAGAAATTTGAATTAAGAAAACATATGATACCTTCAGGCCCATCTTCACTCCCTGT 120
DB 61 CAGAAAGAAATTTGAATTAAGAAAACATATGATACCTTCAGGCCCATCTTCACTCCCTGT 120
QY 121 CTTCTTATGCTTTATTTGGCACTGATGATGAGCCCAAGGGGAAAGTTTAACTGACCCCTG 180
DB 121 CTTCTTATGCTTTATTTGGCACTGATGATGAGCCCAAGGGGAAAGTTTAACTGACCCCTG 180
QY 181 AAACCCATGACATTTTCTATTTATGAGGCCAAGAACCCGATCAATTAATTCAGATT 240
DB 181 AAACCCATGACATTTTCTATTTATGAGGCCAAGAACCCGATCAATTAATTCAGATT 240
QY 241 AAGGCCAATCTCTGCTGTGACTTTTGAACCTAACTGGGAGACAGACAATATTGTG 300
DB 241 AAGGCCAATCTCTGCTGTGACTTTTGAACCTAACTGGGAGACAGACAATATTGTG 300
QY 301 ATAGAACGGAGGAGACTTCTGTATTACACAGAGCCCTTGACAGGAAACAAAGATCTACT 360
DB 301 ATAGAACGGAGGAGACTTCTGTATTACACAGAGCCCTTGACAGGAAACAAAGATCTACT 360
QY 361 CACAAATCTCCAGGTGTGAGCCCTGAGAGCTTAATGAAATTATGAGGGGTCCAGTCCCT 420
DB 361 CACAAATCTCCAGGTGTGAGCCCTGAGAGCTTAATGAAATTATGAGGGGTCCAGTCCCT 420
QY 421 ATCACAATAGAGTGAAGACATCAACAGACATGACCCAGTTTCTCCAGTCAAAAGTAC 480
DB 421 ATCACAATAGAGTGAAGACATCAACAGACATGACCCAGTTTCTCCAGTCAAAAGTAC 480
QY 481 GAAAGGCTCAGTAAAGGCAAGCTCTCGCCAGAGAAAGCCCTTCTGTATGTCAATGACACA 540
DB 481 GAAAGGCTCAGTAAAGGCAAGCTCTCGCCAGAGAAAGCCCTTCTGTATGTCAATGACACA 540
QY 541 GACCTGATGATATCCGGCCATCTCCCAATGGCCAGCTTTATTAACAATGTCTATCCAGCTT 600
DB 541 GACCTGATGATATCCGGCCATCTCCCAATGGCCAGCTTTATTAACAATGTCTATCCAGCTT 600
QY 601 CCCATGATCAACAATGTCTATGATCTTTCAATCAACAACAAAGGGAGCATCTCTCTT 660
DB 601 CCCATGATCAACAATGTCTATGATCTTTCAATCAACAACAAAGGGAGCATCTCTCTT 660
QY 661 ACCCGAGAGGAGATCTCAGGAATTTGAATCTCTTAAGAAATCTTCTATATCTGTGATC 720

```

```

DB 661 ACCCGAGAGGAGATCTCAGGAATTTGAATCTCTTAAGAAATCTTCTATATCTGTGATC 720
QY 721 TCAGTGAAGACATGAGGAGCCAGAGTGAAGAAATTCCTTCAGTATATCCACATCTGTGAT 780
DB 721 TCAGTGAAGACATGAGGAGCCAGAGTGAAGAAATTCCTTCAGTATATCCACATCTGTGAT 780
QY 781 ATCATAGTGAACAGAAATATTTGAAAGACACCAAAACCTGTGAGATGTGGAAGAACTCA 840
DB 781 ATCATAGTGAACAGAAATATTTGAAAGACACCAAAACCTGTGAGATGTGGAAGAACTCA 840
QY 841 ACTGATCTCAACCCCATCAAAATCACTCAGTGCCTGGAATGATCCCGGTGACAAATAT 900
DB 841 ACTGATCTCAACCCCATCAAAATCACTCAGTGCCTGGAATGATCCCGGTGACAAATAT 900
QY 901 TCCTTATGTTGACAAAGAGAGCTGCAAGATTCCTTCATTTGCAATTTGACAGAGAT 960
DB 901 TCCTTATGTTGACAAAGAGAGCTGCAAGATTCCTTCATTTGCAATTTGACAGAGAT 960
QY 961 ATTTACGTGACTCAGCCCTTGGACCGAGAAAGAAAGATGATATTTTATGCAATT 1020
DB 961 ATTTACGTGACTCAGCCCTTGGACCGAGAAAGAAAGATGATATTTTATGCAATT 1020
QY 1021 GCAAAGATGATGAGGAAAGACCACTTTCATATCCGCTGGAATTCATGTAAAGTTTAA 1080
DB 1021 GCAAAGATGATGAGGAAAGACCACTTTCATATCCGCTGGAATTCATGTAAAGTTTAA 1080
QY 1081 GATATTATGATATATCACTACATGTCCTGACACGATACCGTATTTGAGTTCAGAG 1140
DB 1081 GATATTATGATATATCACTACATGTCCTGACACGATACCGTATTTGAGTTCAGAG 1140
QY 1141 AATGAACGACTGGGTAAAGATGAGGAAACCTTACGACATGACAGGAGTGAAGAAAT 1200
DB 1141 AATGAACGACTGGGTAAAGATGAGGAAACCTTACGACATGACAGGAGTGAAGAAAT 1200
QY 1201 ACTGCCAAGATTTTCTAAACCTACAGATTTGAGACAAACCTCCAACTTCCCATGAT 1260
DB 1201 ACTGCCAAGATTTTCTAAACCTACAGATTTGAGACAAACCTCCAACTTCCCATGAT 1260
QY 1261 GGAATCTTCTTAATCCAAACCTATGCTGATATCTACAGTTAGCTTAAACAGTCTTGAAG 1320
DB 1261 GGAATCTTCTTAATCCAAACCTATGCTGATATCTACAGTTAGCTTAAACAGTCTTGAAG 1320
QY 1321 AAGCAAGATATCTCTAGTAAACCTTAAGATGAGGTGTCGACAAAGATTTCAAGACC 1380
DB 1321 AAGCAAGATATCTCTAGTAAACCTTAAGATGAGGTGTCGACAAAGATTTCAAGACC 1380
QY 1381 CTTTGTGTTGTGCAATCAACGTTATGATATCAATGATCCCATCTTTGAAAAA 1440
DB 1381 CTTTGTGTTGTGCAATCAACGTTATGATATCAATGATCCCATCTTTGAAAAA 1440
QY 1441 TCAGATTATGAGAAACCTGACTCTTGTGTAAGACAAACATTTGGTCCACCATTTAAC 1500
DB 1441 TCAGATTATGAGAAACCTGACTCTTGTGTAAGACAAACATTTGGTCCACCATTTAAC 1500
QY 1501 ATCCAGGCACTGATGCTATGATAGCCATTACTGGGAGTTCTTAAATTTCTGATCTATTC 1560
DB 1501 ATCCAGGCACTGATGCTATGATAGCCATTACTGGGAGTTCTTAAATTTCTGATCTATTC 1560
QY 1561 ATTAAGGAGAGAGTGAAGGAGCGCTGGGGTTGACACAGATCCCATACCAACCGGA 1620
DB 1561 ATTAAGGAGAGAGTGAAGGAGCGCTGGGGTTGACACAGATCCCATACCAACCGGA 1620
QY 1621 TATGTCAATATTAAGAGCTCTTGAATTTTGAACAGAGCTGTTTCAACATGTGTTC 1680
DB 1621 TATGTCAATATTAAGAGCTCTTGAATTTTGAACAGAGCTGTTTCAACATGTGTTC 1680
QY 1681 AAAGAGAGAAATCCGAGCTCTAGGTTTGTGTAAGATTAACAGCAAGTTCTTTGGC 1740
DB 1681 AAAGAGAGAAATCCGAGCTCTAGGTTTGTGTAAGATTAACAGCAAGTTCTTTGGC 1740
QY 1741 AAGTTCACGCTTATTTGTGACAGATGATGAAGACCTCAATTTTCCCAACGATATTC 1800
DB 1741 AAGTTCACGCTTATTTGTGACAGATGATGAAGACCTCAATTTTCCCAACGATATTC 1800

```


Db 1741 AAGTCAAGCTATTGTGACAGATGTGATGAAGACCTCAATTTTCCCAACGATATC 1800
Qy 1801 CAAGCGAAAGTCAGTGAAGATGTAGCTATAGGCACTAAAGTGGCAATGTGACTGCCAAG 1860
Db 1801 CAAGCGAAAGTCAGTGAAGATGTAGCTATAGGCACTAAAGTGGCAATGTGACTGCCAAG 1860
Qy 1861 GATCCGAAGGCTGCGACATTAAGCTATTCATGAGGGGAGACACAAAGGTGGCTTAA 1920
Db 1861 GATCCGAAGGCTGCGACATTAAGCTATTCATGAGGGGAGACACAAAGGTGGCTTAA 1920
Qy 1921 ATTGACCAAGCTGACGTGGTGAAGATCTTATAGTGGCTCATTTGACAGAGAAGCCGAAGT 1980
Db 1921 ATTGACCAAGCTGACGTGGTGAAGATCTTATAGTGGCTCATTTGACAGAGAAGCCGAAGT 1980
Qy 1981 CCATATCGGGTACAGTGTGGCCACAGAAAGTGGGGGCTTCTTAAAGTCTGTGTCA 2040
Db 1981 CCATATCGGGTACAGTGTGGCCACAGAAAGTGGGGGCTTCTTAAAGTCTGTGTCA 2040
Qy 2041 GAGTTCAAGCTGATCTTATGATGATGAATGACAAACCTCCAGGGCTAGCCAGAGACTAC 2100
Db 2041 GAGTTCAAGCTGATCTTATGATGATGAATGACAAACCTCCAGGGCTAGCCAGAGACTAC 2100
Qy 2101 ACGGGCTTGTCTTCTCCATCCCTCAGTGCACCTGGAAGTCTCATTTTTCGAGGCTACT 2160
Db 2101 ACGGGCTTGTCTTCTCCATCCCTCAGTGCACCTGGAAGTCTCATTTTTCGAGGCTACT 2160
Qy 2161 GATGATGATGACACTTATTTCCGGGGTCCCATTTTACATTTTCCCTGGCAGTGGAAAC 2220
Db 2161 GATGATGATGACACTTATTTCCGGGGTCCCATTTTACATTTTCCCTGGCAGTGGAAAC 2220
Qy 2221 TTACAAAACGCTGGAAAGTTTCCAAAATCAATGATGCTCAATGCGGAGTGTACACAGG 2280
Db 2221 TTACAAAACGCTGGAAAGTTTCCAAAATCAATGATGCTCAATGCGGAGTGTACACAGG 2280
Qy 2281 CACACAGACTTTGAGAGAGGCGTATGCTGCTTGAATCCGATCAATGATGGGGGTCCG 2340
Db 2281 CACACAGACTTTGAGAGAGGCGTATGCTGCTTGAATCCGATCAATGATGGGGGTCCG 2340
Qy 2341 CCACCTTGGAGGCAATTTCTTTTACAGATTCATTTGCAATTTGTGTGGAAAGT 2400
Db 2341 CCACCTTGGAGGCAATTTCTTTTACAGATTCATTTGCAATTTGTGTGGAAAGT 2400
Qy 2401 TGTTCGGGCGAGAGGTCACAGACTGGGATCCACTGGGCGATGGCAGTTGTGATA 2460
Db 2401 TGTTCGGGCGAGAGGTCACAGACTGGGATCCACTGGGCGATGGCAGTTGTGATA 2460
Qy 2461 CTGCTGACCAACCTTCTGTGATGTGATAATTTTGAAGAGTGTGTATCCGATTAAG 2520
Db 2461 CTGCTGACCAACCTTCTGTGATGTGATAATTTTGAAGAGTGTGTATCCGATTAAG 2520
Qy 2521 AAGGATAAGGCAAGATATGTTGAAGTGTCTCAAGCATCTGAAGTCAAACTCTGAGA 2580
Db 2521 AAGGATAAGGCAAGATATGTTGAAGTGTCTCAAGCATCTGAAGTCAAACTCTGAGA 2580
Qy 2581 AGCTGAATTGAAAGGAATGTTGAATTTATATAGCAAGTGTATTTACAGCAACACCA 2640
Db 2581 AGCTGAATTGAAAGGAATGTTGAATTTATATAGCAAGTGTATTTACAGCAACACCA 2640
Qy 2641 TCTCATCTTATTTTCACTTTCATCTAAGTGTATTAATTTTAAAGAGATTTCCCTCT 2700
Db 2641 TCTCATCTTATTTTCACTTTCATCTAAGTGTATTAATTTTAAAGAGATTTCCCTCT 2700
Qy 2701 TGTCTTATTAATTTTGTAAATTTTCTTTTGTGAGTGTGTGTGTGTGTGTGTGTGTGT 2760
Db 2701 TGTCTTATTAATTTTGTAAATTTTCTTTTGTGAGTGTGTGTGTGTGTGTGTGTGTGT 2760
Qy 2761 GCTGAGAGACAGT 2820
Db 2761 GCTGAGAGACAGT 2820
Qy 2821 TTCTCTGCTCAGCTTCTCTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2880
Db 2821 TTCTCTGCTCAGCTTCTCTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2880

Qy 2881 AATTTTGTATTTTATATAGACAGGGGTTTGCCATTTGGCCAGGCTGGTCTTGAATC 2940
Db 2881 AATTTTGTATTTTATATAGACAGGGGTTTGCCATTTGGCCAGGCTGGTCTTGAATC 2940
Qy 2941 CTGACGTCAAGTATCTGCTGCTTGTCTGCCAATACAGGATGAACCACTGACCCCA 3000
Db 2941 CTGACGTCAAGTATCTGCTGCTTGTCTGCCAATACAGGATGAACCACTGACCCCA 3000
Qy 3001 CCTACTTAATATTTATGATGTGTATAGACATTTAGAGATTTTTCATTTTCCATACAT 3060
Db 3001 CCTACTTAATATTTATGATGTGTATAGACATTTAGAGATTTTTCATTTTCCATACAT 3060
Qy 3061 TTTTCCCTCTGCAAAATGAGCTTACTGTTGTTTCCCTTTGGGGCAAGACACT 3120
Db 3061 TTTTCCCTCTGCAAAATGAGCTTACTGTTGTTTCCCTTTGGGGCAAGACACT 3120
Qy 3121 CATTAATATTTCTGATCAATTTTCTTTTACAGAGATATATCAAGTGTGTCTCATAGA 3180
Db 3121 CATTAATATTTCTGATCAATTTTCTTTTACAGAGATATATCAAGTGTGTCTCATAGA 3180
Qy 3181 ACTGCTGATTCATTTATGTTTCTGATTCATCCGTGTCCCTTCATCTTGCAC 3240
Db 3181 ACTGCTGATTCATTTATGTTTCTGATTCATCCGTGTCCCTTCATCTTGCAC 3240
Qy 3241 TCCCTTGTATTTTCACTGAATTTCAAACTTTGTGAGAGAAAGAAAGTGAAGTCTCAG 3300
Db 3241 TCCCTTGTATTTTCACTGAATTTCAAACTTTGTGAGAGAAAGAAAGTGAAGTCTCAG 3300
Qy 3301 GAAAAATTAATTAATAAGAACAGCCTTTTGGCGCGGCAATTC 3345
Db 3301 GAAAAATTAATTAATAAGAACAGCCTTTTGGCGCGGCAATTC 3345

RESULT 5
US-09-880-107-3288
; Sequence 3288, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3288
; LENGTH: 3345
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U07969
US-09-880-107-3288

Query Match 100.0%; Score 3345; DB 10; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAATTCGCTCGACCACTGAATGAGAAAGACTTTTAAACCAATTTGTGACTTA 60
Db 1 GAATTCGCTCGACCACTGAATGAGAAAGACTTTTAAACCAATTTGTGACTTA 60
Qy 61 CAGAAAGAAATTTGAATTAAGAAATCTATATCTTACAGCCCATCTTCACTCCCTGTGT 120
Db 61 CAGAAAGAAATTTGAATTAAGAAATCTATATCTTACAGCCCATCTTCACTCCCTGTGT 120

QY 121 CTTCTTATGCTTTATTTGGCACTGATATGCGCAAGAGGGAGTTAGTGAAGCCCTG 180
Db 121 CTTCTTATGCTTTATTTGGCACTGATATGCGCAAGAGGGAGTTAGTGAAGCCCTG 180
QY 181 AAACCCATGACATTTCTATTTATGAAGGCAAGACCGAGTCAATTAATTTCCAGTTT 240
Db 181 AAACCCATGACATTTCTATTTATGAAGGCAAGACCGAGTCAATTAATTTCCAGTTT 240
QY 241 AAGGCCAATCCTCTGCTGTGATCTTTGAACTTAAGTGGGAGACAGCAATATTTGTG 300
Db 241 AAGGCCAATCCTCTGCTGTGATCTTTGAACTTAAGTGGGAGACAGCAATATTTGTG 300
QY 301 ATGAAGCGGAGGAGCTTTCTGTATTAACAAGACCTTTGACAGGAGAAACAAGTCTACT 360
Db 301 ATGAAGCGGAGGAGCTTTCTGTATTAACAAGACCTTTGACAGGAGAAACAAGTCTACT 360
QY 361 CACAAATTCAGGTTTGAGCCCTGGAGCGTTAATGGAATTAATGAGGAGGCTCCAGTCCCT 420
Db 361 CACAAATTCAGGTTTGAGCCCTGGAGCGTTAATGGAATTAATGAGGAGGCTCCAGTCCCT 420
QY 421 ATCAACATAGAAGTGAAGACATCAAGCAACATGACCCAGTTTCTCCAGTCAAGTATAC 480
Db 421 ATCAACATAGAAGTGAAGACATCAAGCAACATGACCCAGTTTCTCCAGTCAAGTATAC 480
QY 481 GAAAGCTCAGTAAAGGAGAACTCTGCTGCGCCAGAGAAAGCCCTTCTGTATGTCATGCCACA 540
Db 481 GAAAGCTCAGTAAAGGAGAACTCTGCTGCGCCAGAGAAAGCCCTTCTGTATGTCATGCCACA 540
QY 541 GACCTGATATATCCGGGCACTCCCAATGGCCAGTTTATTAACAGATTTGTCATCCAGTT 600
Db 541 GACCTGATATATCCGGGCACTCCCAATGGCCAGTTTATTAACAGATTTGTCATCCAGTT 600
QY 601 CCCATGATCAACATGTCATGTAATCTTCAATCAACAAACGAGGAGCATCTCTT 660
Db 601 CCCATGATCAACATGTCATGTAATCTTCAATCAACAAACGAGGAGCATCTCTT 660
QY 661 ACCCGAGAGGAGATCTCAGAAATGGAATCTGCTAAGAATCTTCTTAATTCGTGTATC 720
Db 661 ACCCGAGAGGAGATCTCAGAAATGGAATCTGCTAAGAATCTTCTTAATTCGTGTATC 720
QY 721 TCAGTGAAGACATGGGAGGCGCAGAGTGAATTCCTTCAGTATATCAATCTGTGTGAT 780
Db 721 TCAGTGAAGACATGGGAGGCGCAGAGTGAATTCCTTCAGTATATCAATCTGTGTGAT 780
QY 781 ATCATAGTGCAGAGAAATATTTGGAAAGCAACAACCTGTGGAGATGAGGAAATCTCA 840
Db 781 ATCATAGTGCAGAGAAATATTTGGAAAGCAACAACCTGTGGAGATGAGGAAATCTCA 840
QY 841 ACTGATCTCAGCCCATCAAAATCACTCAGGTGCGGTGGAATGATCCCGGTGCAATAT 900
Db 841 ACTGATCTCAGCCCATCAAAATCACTCAGGTGCGGTGGAATGATCCCGGTGCAATAT 900
QY 901 TCCCTTATGTTGACAAAGAGAGCTGCGCAAGTTCCTCATTTTCAATTGACAGAAAGAT 960
Db 901 TCCCTTATGTTGACAAAGAGAGCTGCGCAAGTTCCTCATTTTCAATTGACAGAAAGAT 960
QY 961 ATTATAGTGCATGAGCCCTTGAACCGAGAGAAAGATGATATGTTTTTATGCACTT 1020
Db 961 ATTATAGTGCATGAGCCCTTGAACCGAGAGAAAGATGATATGTTTTTATGCACTT 1020
QY 1021 GCAAAGATAGTACGAGAAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTAA 1080
Db 1021 GCAAAGATAGTACGAGAAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTAA 1080
QY 1081 GATATTAATGAATATCACTACATATGTCGTGACAGTAAACGTTATTTGAGGTCCAGAG 1140
Db 1081 GATATTAATGAATATCACTACATATGTCGTGACAGTAAACGTTATTTGAGGTCCAGAG 1140
QY 1141 AATGAAGACTGGGTAAACATATCGGAGACCTTATCTGACATGACAGGAGTAAAGAAAT 1200
Db 1141 AATGAAGACTGGGTAAACATATCGGAGACCTTATCTGACATGACAGGAGTAAAGAAAT 1200
QY 1201 ACTGCCAAAGTTTTCTAACTACAGATTTGTGAGCAAACTCCCAAATCTTCCATGAT 1260

Db 1201 ACTGCCAAAGTTTTCTAACTACAGATTTGTGAGCAAACTCCCAAATCTTCCATGAT 1260
QY 1261 GGAATCTTCTAATCCAAACCTATGCTGGAATGTTACAGTTACTTAAACGTCCTTAAG 1320
Db 1261 GGAATCTTCTAATCCAAACCTATGCTGGAATGTTACAGTTACTTAAACGTCCTTAAG 1320
QY 1321 AAGCAAGATATCCTCCTGATGCAACTTAACGATAGAGGTGCTGACAAAGATTTCAAGACC 1380
Db 1321 AAGCAAGATATCCTCCTGATGCAACTTAACGATAGAGGTGCTGACAAAGATTTCAAGACC 1380
QY 1381 CTTTGTGTTGTGCAAAATCAACGTTATGATATCAATGATCAGATCCCATCTTTGAAAAA 1440
Db 1381 CTTTGTGTTGTGCAAAATCAACGTTATGATATCAATGATCAGATCCCATCTTTGAAAAA 1440
QY 1441 TCAGATTAATGGAACCTGATCTTCTGCTGAAAGACAAACCTTGGGTCCACATCTTAACC 1500
Db 1441 TCAGATTAATGGAACCTGATCTTCTGCTGAAAGACAAACCTTGGGTCCACATCTTAACC 1500
QY 1501 ATCCAGGCCATGATGCTGATGAGCCATTTACTGGGAGTTCTAATAATCTGTATCATATC 1560
Db 1501 ATCCAGGCCATGATGCTGATGAGCCATTTACTGGGAGTTCTAATAATCTGTATCATATC 1560
QY 1561 ATAAAGGAGACAGTGAAGGAGCGCTGGGGGTTGAACAAGATCCCATACCAACCCGGA 1620
Db 1561 ATAAAGGAGACAGTGAAGGAGCGCTGGGGGTTGAACAAGATCCCATACCAACCCGGA 1620
QY 1621 TATGTCAATATTTAAAAAGCTCTTGTATTTGAAACAGCAGCTGTTTCCAACTGTGTTCC 1680
Db 1621 TATGTCAATATTTAAAAAGCTCTTGTATTTGAAACAGCAGCTGTTTCCAACTGTGTTCC 1680
QY 1681 AAAGCAGAAAAATCCTGAGCCTCTAGTGTGTTGATGTAAGTAAGTCAATGCAAGTCTTTTGGC 1740
Db 1681 AAAGCAGAAAAATCCTGAGCCTCTAGTGTGTTGATGTAAGTAAGTCAATGCAAGTCTTTTGGC 1740
QY 1741 AAGTTCACGCTTATTTGTGACAGATGTAAGTAAGTCAATGTTTCCCAACAGTATTC 1800
Db 1741 AAGTTCACGCTTATTTGTGACAGATGTAAGTAAGTCAATGTTTCCCAACAGTATTC 1800
QY 1801 CAAGCCAAAGTCAATGAGATGATGCTATAGGCACTTAAGTGGCAATGTGATGCCAAG 1860
Db 1801 CAAGCCAAAGTCAATGAGATGATGCTATAGGCACTTAAGTGGCAATGTGATGCCAAG 1860
QY 1861 GATCCAGAAAGTGTGACATATAGCTATTCATGAGGGGAGACACAAGAGGTTGGCTTAA 1920
Db 1861 GATCCAGAAAGTGTGACATATAGCTATTCATGAGGGGAGACACAAGAGGTTGGCTTAA 1920
QY 1921 ATTGACACGATGATGATGATCTTTAGTGTGCTCAATTGACAGAGAACCGGAAGT 1980
Db 1921 ATTGACACGATGATGATGATCTTTAGTGTGCTCAATTGACAGAGAACCGGAAGT 1980
QY 1981 CCATATCGGGTACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
Db 1981 CCATATCGGGTACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
QY 2041 GAGTTCACCTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Db 2041 GAGTTCACCTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
QY 2101 ACGGGCTTGTCTTCCGATCCCTCAGTGCACCTGGAAGTCTCATTTTTCAGAGCTACT 2160
Db 2101 ACGGGCTTGTCTTCCGATCCCTCAGTGCACCTGGAAGTCTCATTTTTCAGAGCTACT 2160
QY 2161 GATGATGATCAGATCTTATTTGGGGGTCCCATTTTCAATTTTCCCTGGCAGTGAAGC 2220
Db 2161 GATGATGATCAGATCTTATTTGGGGGTCCCATTTTCAATTTTCCCTGGCAGTGAAGC 2220
QY 2221 TTAACAACAAGCTGGGAAGTTTCAAAATCAATGATGATGATGATGATGATGATGATGATGAT 2280
Db 2221 TTAACAACAAGCTGGGAAGTTTCAAAATCAATGATGATGATGATGATGATGATGATGATGAT 2280
QY 2281 CACACAGACTTTGAGGAGAGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340

D	2281	CACACACACTTTGAGAGAGAGGCGGTATGTGCTTAAATCCGCATCAATGATGGGGGTCCG	2340
Q	2341	CCACCCCTTGGAGGCACTTGTTCCTTTACCACTTACACTTCTGCACTGTGTGGAAAGAACT	2400
D	2341	CCACCCCTTGGAGGCACTTGTTCCTTTACCACTTACACTTCTGCACTGTGTGGAAAGAACT	2400
Q	2401	TGTTCCGGCCAGAGGACACACACCTGGGATACCACTGGGGCATGGCAGTTGGATATA	2460
D	2401	TGTTCCGGCCAGAGGACACACACCTGGGATACCACTGGGGCATGGCAGTTGGATATA	2460
Q	2461	CTGCTGACCAACCTTCGTGTGATGGATATATTTTAAAGCATTTGTATTCCGCATTAAG	2520
D	2461	CTGCTGACCAACCTTCGTGTGATGGATATATTTTAAAGCATTTGTATTCCGCATTAAG	2520
Q	2521	AAGGATAAAGGCAAAAGATATGTTGAAAGTGCTCAAGCATCTGAAAGTCAAACTCTGAGA	2580
D	2521	AAGGATAAAGGCAAAAGATATGTTGAAAGTGCTCAAGCATCTGAAAGTCAAACTCTGAGA	2580
Q	2581	AGCTGAATTTGAAAAGGAAATGTTGAATTTATATAGCAAGGCAATTTGAGCAACAACA	2640
D	2581	AGCTGAATTTGAAAAGGAAATGTTGAATTTATATAGCAAGGCAATTTGAGCAACAACA	2640
Q	2641	TCTCATCTTATTAATCTTTTCACTCAACGTCATTAATTTTAAACAGATATTCCTCT	2700
D	2641	TCTCATCTTATTAATCTTTTCACTCAACGTCATTAATTTTAAACAGATATTCCTCT	2700
Q	2701	TGTCCTTTAATATTTGCTAAATATTTCTTTTAAAGTGAGTCTGTGTGCCAG	2760
D	2701	TGTCCTTTAATATTTGCTAAATATTTCTTTTAAAGTGAGTCTGTGTGCCAG	2760
Q	2761	GCTGAGATACAGTGTGTGATCCCAAGTCACATGCAACCTCGGCTCCGGGTTCAATGA	2820
D	2761	GCTGAGATACAGTGTGTGATCCCAAGTCACATGCAACCTCGGCTCCGGGTTCAATGA	2820
Q	2821	TTCTCCCTCCAGCTTCTCTAAGTAGCTGGGTTTACAGGCAACCAACCATGCCCCAGCT	2880
D	2821	TTCTCCCTCCAGCTTCTCTAAGTAGCTGGGTTTACAGGCAACCAACCATGCCCCAGCT	2880
Q	2881	AATTTTGTATTTTAAATATAGAGACGGGGTTCGCCATTGSGCCAGAGCTGTCTTGAATCTC	2940
D	2881	AATTTTGTATTTTAAATATAGAGACGGGGTTCGCCATTGSGCCAGAGCTGTCTTGAATCTC	2940
Q	2941	CTGACGTCAAGTGAATCTGCTGCTGGTGTCTCCCAATACAGGCATGAACCATGCAACCA	3000
D	2941	CTGACGTCAAGTGAATCTGCTGCTGGTGTCTCCCAATACAGGCATGAACCATGCAACCA	3000
Q	3001	CCCTACTGATATTTCATGTGCTATAGACATTAGAGAGATTTTTCATTTTCCATGACAT	3060
D	3001	CCCTACTGATATTTCATGTGCTATAGACATTAGAGAGATTTTTCATTTTCCATGACAT	3060
Q	3061	TTTTCCCTCTGCAAAATGGCTTAAGCTACTGTGTTCCTTTTGGGGCAAGACAGACT	3120
D	3061	TTTTCCCTCTGCAAAATGGCTTAAGCTACTGTGTTCCTTTTGGGGCAAGACAGACT	3120
Q	3121	CATTAAATATCTGTACATTTTTCCTTATCAAGAGATATATCATGATGTGTCTCATAGA	3180
D	3121	CATTAAATATCTGTACATTTTTCCTTATCAAGAGATATATCATGATGTGTCTCATAGA	3180
Q	3181	ACTGCTGAGATTCATTTATGTTTTTTCGATTCATCTGCTGTGCTCCCTTCACTTAC	3240
D	3181	ACTGCTGAGATTCATTTATGTTTTTTCGATTCATCTGCTGTGCTCCCTTCACTTAC	3240
Q	3241	TCTTTGGTATTTCACTGATTTTCAAACTTTTGCAGAGAGAAAAGTGAAGACTCAG	3300
D	3241	TCTTTGGTATTTCACTGATTTTCAAACTTTTGCAGAGAGAAAAGTGAAGACTCAG	3300
Q	3301	GAATAATTAATTAATTAAGAAACAGCCTTTTGGCGCGGGAATTC	3345
D	3301	GAATAATTAATTAATTAAGAAACAGCCTTTTGGCGCGGGAATTC	3345

```

1      Patent 118, Application no./09962832
2      Sequence No. US20020110821A1
3      GENERAL INFORMATION:
4      APPLICANT: Ebner, Reinhard
5      TITLE OF INVENTION: Cancer Gene Detect
6      TITLE OF INVENTION: Sets
7      FILE REFERENCE: 689290-74
8      CURRENT APPLICATION NUMBER: US/09/96
9      CURRENT FILING DATE: 2001-09-25
10     PRIOR APPLICATION NUMBER: US/60/235,
11     PRIOR FILING DATE: 2000-09-25
12     PRIOR APPLICATION NUMBER: US/60/235,
13     PRIOR FILING DATE: 2000-09-25
14     NUMBER OF SEQ ID NOS: 259
15     SOFTWARE: PatentIn version 3.0.
16     SEQ ID NO 118
17     LENGTH: 31654
18     TYPE: DNA
19     ORGANISM: Homo sapiens
20     US-09-962-832-118

```

Query Match	99.1%	Score 3314.8	DB 10	Length 3654
Best Local Similarity	99.8%	Pred. No. 0		
Best Match 3319, Conservative	0	Mismatches 7	Indels 0	Gaps 0

QY	8	GTCTGCACCACTGGAATGGGAAGAAAAGACCTTTTAAACACCACTTTTGTGACTTACAGAAAG	67
Db	13	GTCTGCACCACTGGAATGGGAAGAAAAGACCTTTTAAACCACTTTTGTGACTTACAGAAAG	72
QY	68	GAATTGGAATTAAGAAAACCTATGATTACTTACAGGCCATCTTCACTCCCTGTGTCTTCTTA	127
Db	73	GAATTGGAATTAAGAAAACCTATGATTACTTACAGGCCATCTTCACTCCCTGTGTCTTCTTA	132
QY	128	TGCTTTATTTGGCAACTGGATATGGCCAAAGGGGAAGTTTAGTGGACCCCTGAAACCCA	187
Db	133	TGCTTTATTTGGCAACTGGATATGGCCAAAGGGGAAGTTTAGTGGACCCCTGAAACCCA	192
QY	188	TGACATTTTCTATTTATTAAGAGCCCAAGAACCGAGTCAATTAATATATCCAGTTTAAGGCCA	247
Db	193	TGACATTTTCTATTTATTAAGAGCCCAAGAACCGAGTCAATTAATATATCCAGTTTAAGGCCA	252
QY	248	ATCCCTCCGCTGTGACTTTTGAACCTAACCTGGGGAGACAGACACATATTTGTGATAGAAC	307
Db	253	ATCCCTCCGCTGTGACTTTTGAACCTAACCTGGGGAGACAGACACATATTTGTGATAGAAC	312
QY	308	GGGAGGGGACTTCTGTATTACAACAGAGCCTTGGACAGGGAAACAAAGATCTACTCAATC	367
Db	313	GGGAGGGGACTTCTGTATTACAACAGAGCCTTGGACAGGGAAACAAAGATCTACTCAATC	372
QY	368	TCCAGGTTGACAGCCCTGSAACGCTTAATGGAATTAATGTGAGGGGTCAAGTCCCTATCACCA	427
Db	373	TCCAGGTTGACAGCCCTGSAACGCTTAATGGAATTAATGTGAGGGGTCAAGTCCCTATCACCA	432
QY	428	TGGAATGGAAGGACATCAACGACATTCGACCCAGTTTCTCCAGTCCAAAGTACGAGGCT	487
Db	433	TAAAAGTAAAGGACATCAACGACATTCGACCCAGTTTCTCCAGTCCAAAGTACGAGGCT	492
QY	488	CAGTAAGGACAGAACTCTCGCCAGGAAAGCCCTTTGTATGTCAATGACAGACCTTGG	547
Db	493	CAGTAAGGACAGAACTCTCGCCAGGAAAGCCCTTTGTATGTCAATGACAGACCTTGG	552
QY	548	ATGATCCGGGCACTCCCAATGGCCAGCTTATTAACAAATTGATCATCCAGCTTCCCATGA	607
Db	553	ATGATCCGGGCACTCCCAATGGCCAGCTTATTAACAAATTGATCATCCAGCTTCCCATGA	612
QY	608	TCACCAATGTCAATGACTTTTCAATCAACACAAAACGGAGCCATCTCTTACCCGAG	667
Db	613	TCACCAATGTCAATGACTTTTCAATCAACACAAAACGGAGCCATCTCTTACCCGAG	672
QY	668	AGGAGTTCAGGAATTGAATCCTGTAAGAATCCTTCCATATATCGGATCTCAGTGA	727
Db	673	AGGAGTTCAGGAATTGAATCCTGTAAGAATCCTTCCATATATCGGATCTCAGTGA	732

QY 728 AGGACATGGAGGCCCAAGATGAAATTCCTTCACTGATACCAATCTGTGTATATCATAG 787
DB 733 AGGCATATGGAGGAGCCAGATGAGAAATTCCTTCACTGATACCAATCTGTGTATATCATAG 792
QY 788 TGACAGAAATATTTGGAAAGCACCAAAACCTGTGAGATGTGTGAAAACCTCAACTATC 847
DB 793 TGACAGAAATATTTGGAAAGCACCAAAACCTGTGAGATGTGTGAAAACCTCAACTATC 852
QY 848 CTCACCCCATCAAAATCACTCAGGTGGGTGGAATGATCCCGGTGCAAAATATCTTAG 907
DB 853 CTCACCCCATCAAAATCACTCAGGTGGGTGGAATGATCCCGGTGCAAAATATCTTAG 912
QY 908 TTGCAAAAGAGAGCTGCGCAAGATTTCCATTTTCAATTGACACAGAAAGATATTTACG 967
DB 913 TTGCAAAAGAGAGCTGCGCAAGATTTCCATTTTCAATTGACACAGAAAGATATTTACG 972
QY 968 TGACTCAGCCCTTGGACCGAGAAAGATGGAATGTTTATGATGTTTATGCGATTGCAAG 1027
DB 973 TGACTCAGCCCTTGGACCGAGAAAGATGGAATGTTTATGATGTTTATGCGATTGCAAG 1032
QY 1028 ATGAGTACGAGAAACCACTTTCATATCCGTGAAATTCATGTAAAGTTAAAGATATTA 1087
DB 1033 ATGAGTACGAGAAACCACTTTCATATCCGTGAAATTCATGTAAAGTTAAAGATATTA 1092
QY 1088 ATGATTAATCCACTTACATGTCCTGACCAAGTAAACCGTATTTGAGTCCAGAGATGAA 1147
DB 1093 ATGATTAATCCACTTACATGTCCTGACCAAGTAAACCGTATTTGAGTCCAGAGATGAA 1152
QY 1148 GACTGGGTAAACAGATGGGAGCCCTTACCTGACATGACAGGATGAAAGAAATCTGCCA 1207
DB 1153 GACTGGGTAAACAGATGGGAGCCCTTACCTGACATGACAGGATGAAAGAAATCTGCCA 1212
QY 1208 ACAATTTTCTTAACTACAGAGATGTGTGAGCAAACTCCCAAACTTCCATGATGAACTCT 1267
DB 1213 ACAATTTTCTTAACTACAGAGATGTGTGAGCAAACTCCCAAACTTCCATGATGAACTCT 1272
QY 1268 TCCCTAATCCAAACCTATGCTGGAAGATGTTACAGTTAGCTTAAACAGTCTTGAAGAGCA 1327
DB 1273 TCCCTAATCCAAACCTATGCTGGAAGATGTTACAGTTAGCTTAAACAGTCTTGAAGAGCA 1332
QY 1328 ATATCTCTGATCAATCTTAAACGATGAGTGTCTGCAAAAGATTTCAAGACCTTTGTT 1387
DB 1333 ATATCTCTGATCAATCTTAAACGATGAGTGTCTGCAAAAGATTTCAAGACCTTTGTT 1392
QY 1388 TTGTGCAAAATCAACGTTATGATATCAATGATCAAGATCCCATCTTGAAGAAATCAAGT 1447
DB 1393 TTGTGCAAAATCAACGTTATGATATCAATGATCAAGATCCCATCTTGAAGAAATCAAGT 1452
QY 1448 ATGAAAACCTGACTCTTGTGTAAGACCAAAACATTTGGGTCCACATCTTAAACATCCAG 1507
DB 1453 ATGAAAACCTGACTCTTGTGTAAGACCAAAACATTTGGGTCCACATCTTAAACATCCAG 1512
QY 1508 CCACTATGCTGATGAGCCATTTACTGGAGTTCTTAAATTTCTGTATCATATCAATAAG 1567
DB 1513 CCACTATGCTGATGAGCCATTTACTGGAGTTCTTAAATTTCTGTATCATATCAATAAG 1572
QY 1568 GAGACAGTGAAGGAGCCCTGGGGTTGACACAGATCCCATACCAACACCGATATGTCA 1627
DB 1573 GAGACAGTGAAGGAGCCCTGGGGTTGACACAGATCCCATACCAACACCGATATGTCA 1632
QY 1628 TAAATTAATAAAGCTCTTGTATTTGAAAACAGACAGCTGTTTCCAACTTGTGTGCAAGCA 1687
DB 1633 TAAATTAATAAAGCTCTTGTATTTGAAAACAGACAGCTGTTTCCAACTTGTGTGCAAGCA 1692
QY 1688 AAAATCTGAGCCTTATAGTGTGTGTGTGTAAGTAAATGCAATGCAAGTTCTTTTGCAGTTCA 1747
DB 1693 AAAATCTGAGCCTTATAGTGTGTGTGTGTAAGTAAATGCAATGCAAGTTCTTTTGCAGTTCA 1752
QY 1748 CGCTTATTTGTGACAGATGTAATGTAAGCACTCAATTTTCCAAACGTTATCCAGGCA 1807
DB 1753 CGCTTATTTGTGACAGATGTAATGTAAGCACTCAATTTTCCAAACGTTATCCAGGCA 1812
QY 1808 AAGTCAGTGAGATGATGATTATAGCACTTAAAGTGGCAATGTGACTGCCAAGATCCAG 1867

DB 1813 AAGTCAGTGAGATGATGATTATAGCACTTAAAGTGGCAATGTGACTGCCAAGATCCAG 1872
QY 1868 AAGTCTGACATTAAGCTATTTCACTGAGGGGACACCAAGTGGCTTAAATTAACC 1927
DB 1873 AAGTCTGACATTAAGCTATTTCACTGAGGGGACACCAAGTGGCTTAAATTAACC 1932
QY 1928 ACCGTACTGTGAGATCTTATGATGTGCTCATTTGACAGAGAACCGGAATCATATC 1987
DB 1933 ACCGTACTGTGAGATCTTATGATGTGCTCATTTGACAGAGAACCGGAATCATATC 1992
QY 1988 GGGTACAGATGTGGCCACAGAAAGTGGGGGTCTTCTTAAGCTGTGTCAAGATTCC 2047
DB 1993 GGGTACAGATGTGGCCACAGAAAGTGGGGGTCTTCTTAAGCTGTGTCAAGATTCC 2052
QY 2048 ACCGTATCTTATGATGATGTAATGACAAACCTCCAGAGCTAGCCAGAGATCAACGGCT 2107
DB 2053 ACCGTATCTTATGATGATGTAATGACAAACCTCCAGAGCTAGCCAGAGATCAACGGCT 2112
QY 2108 TGTTCCTTGCATATCCCTCAGTGCACCTGGAAGTCTCATTTTCAAGGCTATGATATG 2167
DB 2113 TGTTCCTTGCATATCCCTCAGTGCACCTGGAAGTCTCATTTTCAAGGCTATGATATG 2172
QY 2168 ATGAGCACTTATTTGGGGTCCCATTTTACATTTTCCCTGGCAGTGAAGCTTACAA 2227
DB 2173 ATGAGCACTTATTTGGGGTCCCATTTTACATTTTCCCTGGCAGTGAAGCTTACAA 2232
QY 2228 ACGACTGGGAAGTTTCCAAATCAATGTATCTATCCGACGTCTTACAGGACACAG 2292
DB 2233 ACGACTGGGAAGTTTCCAAATCAATGTATCTATCCGACGTCTTACAGGACACAG 2292
QY 2288 ACTTGAAGAGAGGGGATATGCTGTGATCCGATCAATGATGGGGGTGGCCACCT 2347
DB 2293 ACTTGAAGAGAGGGGATATGCTGTGATCCGATCAATGATGGGGGTGGCCACCT 2352
QY 2348 TGGAGGCAATGTTTCTTTACAGATTACATTTGCAAGTGTGTGGAAGAAAGTTGTTCC 2407
DB 2353 TGGAGGCAATGTTTCTTTACAGATTACATTTGCAAGTGTGTGGAAGAAAGTTGTTCC 2412
QY 2408 GGCACACAGTACACAGACTGGGATACCACTGTGGGCAATGGCAATGGATATCTGTA 2467
DB 2413 GGCACACAGTACACAGACTGGGATACCACTGTGGGCAATGGCAATGGATATCTGTA 2472
QY 2468 CCACCTCTGAGTATGATGATTAATTTAGCAGTGTGTTTACCGCATTAAGAGATA 2527
DB 2473 CCACCTCTGAGTATGATGATTAATTTAGCAGTGTGTTTACCGCATTAAGAGATA 2532
QY 2528 AAGGCAAGATTAATGTTGAAGTCTCAAGCATCTGAACTCAAACTGTGAGAGCTGAA 2587
DB 2533 AAGGCAAGATTAATGTTGAAGTCTCAAGCATCTGAACTCAAACTGTGAGAGCTGAA 2592
QY 2588 TTTGAAAAGAAATGTTGAATTAATTAATGCAAGTGTATTTGACGAAACCACTGTATC 2647
DB 2593 TTTGAAAAGAAATGTTGAATTAATTAATGCAAGTGTATTTGACGAAACCACTGTATC 2652
QY 2648 CTATTACTTTTCACTAAGCTGATATTAATTTTAAACAGATATTCCTCTGTGCTT 2707
DB 2653 CTATTACTTTTCACTAAGCTGATATTAATTTTAAACAGATATTCCTCTGTGCTT 2712
QY 2708 TAAATTTGCTAAATATTTCTTTTGAAGTGAAGTCTTGTCTGTGCGCCAGGCTGAG 2767
DB 2713 TAAATTTGCTAAATATTTCTTTTGAAGTGAAGTCTTGTCTGTGCGCCAGGCTGAG 2772
QY 2768 TACAGTGTGTGATCCCACTGACATCTCCGCTCTGTGGTTCAATGATCTCT 2827
DB 2773 TACAGTGTGTGATCCCACTGACATCTCCGCTCTGTGGTTCAATGATCTCT 2832
QY 2828 GCTCAGCTTCCATAGTATGCTGGTTTACAGGACCCACACATGCCCAGATATTTT 2887
DB 2833 GCTCAGCTTCCATAGTATGCTGGTTTACAGGACCCACACATGCCCAGATATTTT 2892
QY 2888 GTATTTTAAAGAGCGGGGTTTGGCAATTTGGCAGGCTGTGAACTCTGACGT 2947

QY	1266	TCCTATCCAAACCTATGCTGGAAATGTTACAGTATAGCTAAACAGTCTTGAAGAAACAAG	132
Db	1273	TCCTATCCAAACCTATGCTGGAAATGTTACAGTATAGCTAAACAGTCTTGAAGAAACAAG	133
QY	1328	ATACTCTTCAGTACAACTTAAACGATAGAGGTGTCTGACAAAGATTTCAGACCCCTTGT	138
Db	1333	ATACTCTTCAGTACAACTTAAACGATAGAGGTGTCTGACAAAGATTTCAGACCCCTTGT	139
QY	1388	TTGTGCAATCAACGTTATTTGATATCAATGATCCGATCTTTGAAAAATCAGATT	144
Db	1393	TTGTGCAATCAACGTTATTTGATATCAATGATCCGATCTTTGAAAAATCAGATT	145
QY	1448	ATGAAAACCTGACTCTGTGCTGAAGACAAAACATTGGGGTCCGACGATCTTAACATCCAG	150
Db	1453	ATGAAAACCTGACTCTGTGCTGAAGACAAAACATTGGGGTCCGACGATCTTAACATCCAG	151
QY	1508	CCACTGATGCTGATGAGCCATTACTGGAGGTTCTAAAATTTCTGATCAATATCAAAAG	156
Db	1513	CCACTGATGCTGATGAGCCATTACTGGAGGTTCTAAAATTTCTGATCAATATCAAAAG	157
QY	1568	GAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCAACCAACCGGATATGTCA	162
Db	1573	GAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCAACCAACCGGATATGTCA	163
QY	1628	TAAATTAAGGCTCTGTGATTTTGAACACGACGCTTTCCAACTTGTGTCAAGCAG	168
Db	1633	TAAATTAAGGCTCTGTGATTTTGAACACGACGCTTTCCAACTTGTGTCAAGCAG	169
QY	1688	AAATCTGAGCCTCTAGTGTTGGTGTGCAAGTAAATGCAAGTCTTTGGCAAGTTCA	174
Db	1693	AAATCTGAGCCTCTAGTGTTGGTGTGCAAGTAAATGCAAGTCTTTGGCAAGTTCA	175
QY	1748	CGCTTATTTGTGACAGATGTGAATGAACACTCTAAATTTTCCCAACGTAATCCAGCGA	180
Db	1753	CGCTTATTTGTGACAGATGTGAATGAACACTCTAAATTTTCCCAACGTAATCCAGCGA	181
QY	1808	AAGTAGAGGAGATGATGACTAATGACACTAAAGTGGCAATGTGCTGCCAAGATCCAG	186
Db	1813	AAGTAGAGGAGATGATGACTAATGACACTAAAGTGGCAATGTGCTGCCAAGATCCAG	187
QY	1868	AAGTCTGGAACATAAGCTATTCACTGAGGGGAGACAAAGAGTTGGCTTAAATTTGACC	192
Db	1873	AAGTCTGGAACATAAGCTATTCACTGAGGGGAGACAAAGAGTTGGCTTAAATTTGACC	193
QY	1928	ACGTGACTGTGAGATCTTTAGTGTGGCTCCAATTGACAGAGAACCCGGAATCCATATC	198
Db	1933	ACGTGACTGTGAGATCTTTAGTGTGGCTCCAATTGACAGAGAACCCGGAATCCATATC	199
QY	1988	GGGTACAAGTGTGGCCACAAAGATGAGGGGGTCTTCTTAAAGCTCTGTGTCAAGTTCC	204
Db	1993	GGGTACAAGTGTGGCCACAAAGATGAGGGGGTCTTCTTAAAGCTCTGTGTCAAGTTCC	205
QY	2048	ACCTGATCCTTATGAGATGTAATGACAAACCCCTCCAGGCTGAGCCAAAGACTACAGGGCT	210
Db	2053	ACCTGATCCTTATGAGATGTAATGACAAACCCCTCCAGGCTGAGCCAAAGACTACAGGGCT	211
QY	2108	TGTTCTTGTGCATCCCTCAGGTGACCTGGAAGTCTGATTTTCAGGCTACTGATGATG	216
Db	2113	TGTTCTTGTGCATCCCTCAGGTGACCTGGAAGTCTGATTTTCAGGCTACTGATGATG	217
QY	2168	ATCAGACTTATTTGGGGTCCCATTTTACATTTTCCCTGGGCACTGGAAGCTTACAA	222
Db	2173	ATCAGACTTATTTGGGGTCCCATTTTACATTTTCCCTGGGCACTGGAAGCTTACAA	223
QY	2228	ACGACTGGGAAGTTTCCAAATCATATGTAATCATGAGCCGACGTGTCTACAGGACACAG	228
Db	2233	ACGACTGGGAAGTTTCCAAATCATATGTAATCATGAGCCGACGTGTCTACAGGACACAG	229
QY	2288	ACTTTGAGAGAGGGCGTATGTGCTTTGATCCGATCCGATCAATGATGGGGTGGCCACCT	234
Db	2293	ACTTTGAGAGAGGGCGTATGTGCTTTGATCCGATCCGATCAATGATGGGGTGGCCACCT	235
QY	2348	TGGAAGCATTTGTTCTTTACACGTTACATTTCTGCAGTTGTGTGGAAGAAAGTTGTTTC	240

Db	2353	TGGAAAGCATGTGTTCTTTACCACTACATTCTGCAAGTTGTGTGAAGAAAGTTGTTTCC	2412
Qy	2408	GGCCGACAGGCTCACCAAGCTGGGATACCACCTGTGGGCATGCGAGTTGGTATACCTGTGA	2467
Db	2413	GGCCGACGAGCTCACAGACTGGGATACCACCTGTGGGCATGCGAGTTGGTATACCTGTGA	2472
Qy	2468	CCACCCCTCTGSGTATTTGGTATTAATTTTGAAGAGTTGGTTATCCGGATTAAGAAGGATA	2527
Db	2473	CCACCCCTCTGSGTATTTGGTATTAATTTTGAAGAGTTGGTTATCCGGATTAAGAAGGATA	2532
Qy	2528	AAGGCAAAAGATATGTTGAAAAGTGCCTAAGCATCTGAAGTCAAACTCTGAGAAGCTGAA	2587
Db	2533	AAGGCAAAAGATATGTTGAAAAGTGCCTAAGCATCTGAAGTCAAACTCTGAGAAGCTGAA	2592
Qy	2588	TTTGGAAAAGGAATGTTTGAATTTATATAGCAAGTGTATTTCAAGCAACAACATCTCATC	2647
Db	2593	TTTGGAAAAGGAATGTTTGAATTTATATAGCAAGTGTATTTCAAGCAACAACATCTCATC	2652
Qy	2648	CTATTACTCTTTCATCTAAACGTGCATTTAAATTTTAAACAGATATTCCTCTGTGCTT	2707
Db	2653	CTATTACTCTTTCATCTAAACGTGCATTTAAATTTTAAACAGATATTCCTCTGTGCTT	2712
Qy	2708	TAAATATTTGCTAAATATTTCTTTTATAGGTGAGTCTGTGCTGTGCGCCAGGCTGAG	2767
Db	2713	TAAATATTTGCTAAATATTTCTTTTATAGGTGAGTCTGTGCTGTGCGCCAGGCTGAG	2772
Qy	2768	TACAGTGGTGTGATCCCAAGCTCACTGGAACCTCCGCTCCTGGGTTCACATGATCTTCCT	2827
Db	2773	TACAGTGGTGTGATCCCAAGCTCACTGGAACCTCCGCTCCTGGGTTCACATGATCTTCCT	2832
Qy	2828	GCCTCAGGCTTCCTAAGTAGCTGGGTATTAAGGACCCACCAACACATGCCCAGCTAAATTTT	2887
Db	2833	GCCTCAGGCTTCCTAAGTAGCTGGGTATTAAGGACCCACCAACACATGCCCAGCTAAATTTT	2892
Qy	2888	GTAATTTTAAATAGAGAGGGGTTTTCGCCATTTTGGCCAGGCTGTGCTTGAATCCTGACGT	2947
Db	2893	GTAATTTTAAATAGAGAGGGGTTTTCGCCATTTTGGCCAGGCTGTGCTTGAATCCTGACGT	2952
Qy	2948	CAGGTGATCTGCGCTGCGCTTGCTCCCAATTAACAGGACATGAACCACTGACACCACTACTT	3007
Db	2953	CAGGTGATCTGCGCTGCGCTTGCTCCCAATTAACAGGACATGAACCACTGACACCACTACTT	3012
Qy	3008	AGATATTTCATGTGCTATATGACATTTAGAGAGATTTTCAATTTTCCATGACATTTTCT	3067
Db	3013	AGATATTTCATGTGCTATATGACATTTAGAGAGATTTTCAATTTTCCATGACATTTTCT	3072
Qy	3068	CTCTGCAAAATGGCTTACTGACTTGTTGTTTTCCCTTTTGGGGGAAGCAACAACATTTAA	3127
Db	3073	CTCTGCAAAATGGCTTACTGACTTGTTGTTTTCCCTTTTGGGGGAAGCAACAACATTTAA	3132
Qy	3128	TATTCGTGACATTTTTCCTTATCAAGAGATATATACATGTTGTCTCATGAACTGCT	3187
Db	3133	TATTCGTGACATTTTTCCTTATCAAGAGATATATACATGTTGTCTCATGAACTGCT	3192
Qy	3188	GGATTCGATTAATGTTTTCTGATTCGATCCGTGTGCCCCCTTCATCCTTGAATCCTTTG	3247
Db	3193	GGATTCGATTAATGTTTTCTGATTCGATCCGTGTGCCCCCTTCATCCTTGAATCCTTTG	3252
Qy	3248	GTATTTCACTGAATTTCAAAACATTTGTCAAGAGAAAAGTGAAGGACTCAGAGAAAAT	3307
Db	3253	GTATTTCACTGAATTTCAAAACATTTGTCAAGAGAAAAGTGAAGGACTCAGAGAAAAT	3312
Qy	3308	AAATTAATTAAGAAAGCCCTTTTGC 3333	
Db	3313	AAATTAATTAAGAAAGCCCTTTTTC 3338	

RESULT 8
US-10-025-380-1086
; Sequence 1086, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
APPLICANT: Lodes, Michael J.
APPLICANT: Secret, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Stolk, John A.
APPLICANT: Wang, Tongtong
APPLICANT: Jiang, Yuguang
APPLICANT: Smith, Carole L.
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Skeiky, Yasir A. W.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedrick, Thomas S.
APPLICANT: Carter, Derrick
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.471C14
CURRENT APPLICATION NUMBER: US/10/025.380
CURRENT FILING DATE: 2001-12-19
NUMBER OF SEQ ID NOS: 1129
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1086
LENGTH: 2877
TYPE: DNA
ORGANISM: Homo sapiens
US-10-025-380-1086

Query Match 72.7%; Score 2432.6; DB 9; Length 2877;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2432; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 154 CAAGAGGGAAGTTAGTGAACCCCTGAACCCATGACATTTCTATTATGAAGCCAA 213
DB 445 CAAGAGGGAAGTTAGTGAACCCCTGAACCCATGACATTTCTATTATGAAGCCAA 504
QY 214 GAACCGAGTCAATTTATTCAGTTTAAGGCAATCCTCGCTGAGCTTTGAAGCTA 273
DB 505 GAACCGAGTCAATTTATTTCAAGTTTAAGGCAATCCTCGCTGAGCTTTGAAGCTA 564
QY 274 ACTGGGAGAGCAGCAACATATTTGTGTATGAGCGGAGGAGCTTGTATTAACAGAGA 333
DB 565 ACTGGGAGAGCAGCAACATATTTGTGTATGAGCGGAGGAGCTTGTATTAACAGAGA 624
QY 334 GCCTTGAAGAGGAAAGATCTACTCAATCTCCAGGTTGAGCCCTGAGAGCTAAT 393
DB 625 GCCTTGAAGAGGAAAGATCTACTCAATCTCCAGGTTGAGCCCTGAGAGCTAAT 684
QY 394 GGAATTAAGTGAAGGAGGTCAGTCCCTATCCATAGAGTGAAGAGCATCAAGCAAT 453
DB 685 GGAATTAAGTGAAGGAGGTCAGTCCCTATCCATAGAGTGAAGAGCATCAAGCAAT 744
QY 454 CGACCCAGCTTTCTCCAGTCAAGTACAGAGGCTCAGTAAGGAGAACTCTCGCCAGAGA 513
DB 745 CGACCCAGCTTTCTCCAGTCAAGTACAGAGGCTCAGTAAGGAGAACTCTCGCCAGAGA 804
QY 514 AAGCCCTTCTGTATGTCAATGCCAGAGCTGGATGATCCGGCACTCCCAATGGCCAG 573
DB 805 AAGCCCTTCTGTATGTCAATGCCAGAGCTGGATGATCCGGCACTCCCAATGGCCAG 864
QY 574 CTTTATTAACAGATTTGATCAATGCCATGCCATGATCAACATGATGATTTCAATG 633
DB 865 CTTTATTAACAGATTTGATCAATGCCATGCCATGATCAACATGATGATTTCAATG 924
QY 634 AACACAAAAGGAGGAGCATCTCTTACCCGAGAGGAGTCTCAGAAATGATCTGCT 693
DB 925 AACACAAAAGGAGGAGCATCTCTTACCCGAGAGGAGTCTCAGAAATGATCTGCT 984
QY 694 AAGATCTCTTCTATTAATCTGTGATCTCAGTGAAGAGCATGGAGGCGCAGATGAGAT 753
DB 985 AAGATCTCTTCTATTAATCTGTGATCTCAGTGAAGAGCATGGAGGCGCAGATGAGAT 1044

QY 754 TCCTCAGTATACACATCTGTGATATCATAGTACAGAGATATTTGAAAGCACCA 813
DB 1045 TCCTCAGTATACACATCTGTGATATCATAGTACAGAGATATTTGAAAGCACCA 1104
QY 814 AAACCTGTGAGATGTGTGAAAACTCACTGATCTCCACCCATCAAAATCACTCAGGTG 873
DB 1105 AAACCTGTGAGATGTGTGAAAACTCACTGATCTCCACCCATCAAAATCACTCAGGTG 1164
QY 874 CGGTGAAATGATCCGGTGCAGATATTCCTTAGTTGACAAAGAGAGGCTGCCAAGTTG 933
DB 1165 CGGTGAAATGATCCGGTGCAGATATTCCTTAGTTGACAAAGAGAGGCTGCCAAGTTG 1224
QY 934 CCATTTTCAATTGACAGAGAGGATATTTACGTGACTCAGCCCTTGAACCGAGAGAA 993
DB 1225 CCATTTTCAATTGACAGAGAGGATATTTACGTGACTCAGCCCTTGAACCGAGAGAA 1284
QY 994 AAGATGATATGTTTTTATGATGAGTTGCAAGAGATGAGTACGAAAAACCACTTCAAT 1053
DB 1285 AAGATGATATGTTTTTATGATGAGTTGCAAGAGATGAGTACGAAAAACCACTTCAAT 1344
QY 1054 CCGCTGGAATTCATGTAAGTAAAGTTAAGATATTAATGATATCCACTCATGTCGCTCA 1113
DB 1345 CCGCTGGAATTCATGTAAGTAAAGTTAAGATATTAATGATATCCACTCATGTCGCTCA 1404
QY 1114 CCAGTAACCGTATTTGAGTCCAGAGAAATGAACGATGAGTATCGGAGACCTT 1173
DB 1405 CCAGTAACCGTATTTGAGTCCAGAGAAATGAACGATGAGTATCGGAGACCTT 1464
QY 1174 ACTGCATATGACAGGATGAGAAATTAAGTCCCAAGATTTTAACTACAGATTTG 1233
DB 1465 ACTGCATATGACAGGATGAGAAATTAAGTCCCAAGATTTTAACTACAGATTTG 1524
QY 1234 GAGCAAACTCCCAACTCCCATGATGAGAGCTCTTCTTAATCCAAACCTATGCTGAATG 1293
DB 1525 GAGCAAACTCCCAACTCCCATGATGAGAGCTCTTCTTAATCCAAACCTATGCTGAATG 1584
QY 1294 TTACAGTTAGCTAAACAGTCTTGAAGAGCAGATACCTCTCAGTACCACTTAACAGAT 1353
DB 1585 TTACAGTTAGCTAAACAGTCTTGAAGAGCAGATACCTCTCAGTACCACTTAACAGAT 1644
QY 1354 GAGGTGTCTGACAAAGATTTCAAGACCTTTGTTGTGCAATCAAGTTATGATATC 1413
DB 1645 GAGGTGTCTGACAAAGATTTCAAGACCTTTGTTGTGCAATCAAGTTATGATATC 1704
QY 1414 AATGATCAATCCCATCTTTGAAGAAATCAAGATTAAGAACTGCTGTGTAAGAC 1473
DB 1705 AATGATCAATCCCATCTTTGAAGAAATCAAGATTAAGAACTGCTGTGTAAGAC 1764
QY 1474 ACAAACTGGGTCCACATCTTAACATCCAGGCCATGATGATGAGCCATTTACT 1533
DB 1765 ACAAACTGGGTCCACATCTTAACATCCAGGCCATGATGATGAGCCATTTACT 1824
QY 1534 GGGAGTTCTTAAATCTGTATCATATCAATTAAGAGAGACGTAGAGGAGCCCTGGGGGTT 1593
DB 1825 GGGAGTTCTTAAATCTGTATCATATCAATTAAGAGAGACGTAGAGGAGCCCTGGGGGTT 1884
QY 1594 GAGACAGATCCCATCCCAACACCGGATATGATATTAAGAAAGGCTCTGATTTTGA 1653
DB 1885 GAGACAGATCCCATCCCAACACCGGATATGATATTAAGAAAGGCTCTGATTTTGA 1944
QY 1654 ACAGCAGCTGTTTCAACATTTGTTCAAAGAGAGAAATCTGAGCTCTAGTGTGTTG 1713
DB 1945 ACAGCAGCTGTTTCAACATTTGTTCAAAGAGAGAAATCTGAGCTCTAGTGTGTTG 2004
QY 1714 GTGAAGTCAATCAAGATTTCTTTTGCAGAGTCAAGCTTATTTGTGACAGATGTAATGA 1773
DB 2005 GTGAAGTCAATCAAGATTTCTTTTGCAGAGTCAAGCTTATTTGTGACAGATGTAATGA 2064
QY 1774 GCACTTCAATTTTCCCAACGATTTCCAAAGCGAAGTCAAGTGAATGATGATAGG 1833
DB 2065 GCACTTCAATTTTCCCAACGATTTCCAAAGCGAAGTCAAGTGAATGATGATAGG 2124
QY 1834 ACTAAGTGGCAATGATGATCCCAAGAGATCCAGAAAGTCTGAGATTAAGCTATTAAG 1893

Db	2125	ACTTAAAGTGGGACATGTGACTGCAAGATCCAGAAGCTCTGGACATTAAGCATATTCAGTGC	2186
Qy	1894	AGGGGAGACACAAGAGGTTGGCTTAAATTTGACACAGTGACTGGTGAATCTTTAGTGTG	1955
Db	2185	AGGGGAGACACAAGAGGTTGGCTTAAATTTGACACAGTGACTGGTGAATCTTTAGTGTG	2244
Qy	1954	GCTCCATTGGACAGAGAGCCGGAAAGTCCATATCGGGTAAACAAGTGTGGCCACAAACTA	2013
Db	2245	GCTCCATTGGACAGAGAGCCGGAAAGTCCATATCGGGTAAACAAGTGTGGCCACAAACTA	2304
Qy	2014	GGGGGGGTCTTCTTAAAGCTCTGTGTCAAGTTCACAGTTCACATGATCCTTATGGAATGAC	2073
Db	2305	GGGGGGGTCTTCTTAAAGCTCTGTGTCAAGTTCACAGTTCACATGATCCTTATGGAATGAC	2366
Qy	2074	AACTCTCCCAAGGCTAGCCCAAGACTACAAGGGCTTGTCTTTCGCATCCCTCAGTGA	2133
Db	2365	AACTCTCCCAAGGCTAGCCCAAGACTACAAGGGCTTGTCTTTCGCATCCCTCAGTGA	2422
Qy	2134	CCTGGAAGTCTCATTTTGGAGGCTACTGATGATCAAGCATTTATTTCCGGGTCCCAT	2193
Db	2425	CCTGGAAGTCTCATTTTGGAGGCTACTGATGATCAAGCATTTATTTCCGGGTCCCAT	2488
Qy	2194	TTTACATTTTCCCTCGGCGAGTGAAGCTTACAAGACGACTGGGAAGTTCCAAATCAAT	2253
Db	2485	TTTACATTTTCCCTCGGCGAGTGAAGCTTACAAGACGACTGGGAAGTTCCAAATCAAT	2544
Qy	2254	GGTACTCATGCCCCGACTGTCTACACAGGACACAGACTTTGAGAGAGGGCGATATGTCGT	2313
Db	2545	GGTACTCATGCCCCGACTGTCTACACAGGACACAGACTTTGAGAGAGGGCGATATGTCGT	2604
Qy	2314	TTGATCCGCATTAATGATGGGGGTGGGCCACCTTGGAAAGCATTTGTTTACAGTT	2373
Db	2605	TTGATCCGCATTAATGATGGGGGTGGGCCACCTTGGAAAGCATTTGTTTACAGTT	2654
Qy	2374	ACATTCCTGAAGTGTGTGGGAAGAGATTGTTCCGGCCAGACAGTCAACAGCTGGGATA	2433
Db	2655	ACATTCCTGAAGTGTGTGGGAAGAGATTGTTCCGGCCAGACAGTCAACAGCTGGGATA	2724
Qy	2434	CCCACTGTGGGCGATGGCAGTTGGTACTGCTGACCACTTCTGTGATTTGATTAAT	2493
Db	2725	CCCACTGTGGGCGATGGCAGTTGGTACTGCTGACCACTTCTGTGATTTGATTAAT	2784
Qy	2494	TTAGCAGTTGTGTTTATTCGCGATTAAGAAGGCAATTAAGGCAATGTTGAAGTCT	2553
Db	2785	TTAGCAGTTGTGTTTATTCGCGATTAAGAAGGCAATTAAGGCAATGTTGAAGTCT	2844
Qy	2554	CAAGCATCTGAAGTCAAACTCTTGGAAGCTGA	2586
Db	2845	CAAGCATCTGAAGTCAAACTCTTGGAAGCTGA	2877
RESULT 9			
US-09-922-217-1086			
Sequence 1086, Application US/09922217			
Patent No. US2002076414A1			
GENERAL INFORMATION:			
APPLICANT: Xu, Jiangchun			
APPLICANT: Lodes, Michael J.			
APPLICANT: Secrist, Heather			
APPLICANT: Benson, Darin R.			
APPLICANT: Meagher, Madeleine Joy			
APPLICANT: Stolk, John A.			
APPLICANT: Wang, Tongtong			
APPLICANT: Jiang, Yugu			
APPLICANT: Smith, Carole Lynn			
APPLICANT: King, Gordon E.			
APPLICANT: Wang, Aijun			
APPLICANT: Clapper, Jonathan D.			
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS			
FILE REFERENCE: 210121.471C13			
CURRENT APPLICATION NUMBER: US/09/922.217			

Query Match	72.7%; Score 2432.6; DB 10; Length 2877;	
Best Local Similarity	100.0%; Prid. No. 0;	
Matches 2432; Conservative 1; Mismatches 0; Indels 0; Gaps 0;		
Query	154 CAAGAGGGGAGTCTTACGTGACCCCTGTAACCCATGACATTTTCTATTATTGAAAGGCCAA	213
Db	445 CAAGAGGGGAGTCTTACGTGACCCCTGTAACCCATGACATTTTCTATTATTGAAAGGCCAA	504
Query	214 GAACCGAGTCAATTAATTAATTCAGTTTAAAGCCAACTCTGCTGTGACTTTGAACCTA	273
Db	505 GAACCGAGTCAATTAATTAATTCAGTTTAAAGCCAACTCTCTGTGACTTTGAACCTA	564
Query	274 ACTGGGGAGACAGACAAATATTTGTGTATGAAACGGAGGGAATTCTGTATTACAACGA	333
Db	565 ACTGGGGAGACAGACAAATATTTGTGTATGAAACGGAGGGAATTCTGTATTACAACGA	624
Query	334 GCCTTGGACAGGGAACAAAGATCTACTACAACTCCAGGTTGACAGCCCTGGAAGCTAAT	393
Db	625 GCCTTGGACAGGGAACAAAGATCTACTACAACTCCAGGTTGACAGCCCTGGAAGCTAAT	684
Query	394 GGAATTAATGTAGAGGGTCCAGTCCCTATCAACATAGAAAGTGAAGACATCAACGACAA	453
Db	685 GGAATTAATGTAGAGGGTCCAGTCCCTATCAACATAGAAAGTGAAGACATCAACGACAA	744
Query	454 CGACCCAGCTTTCTCCAGTCAAAGTGAAGGCTCAGTAAGGCGAACTTCTGCCACGA	513
Db	745 CGACCCAGCTTTCTCCAGTCAAAGTGAAGGCTCAGTAAGGCGAACTTCTGCCACGA	804
Query	514 AAGCCCTTCTGTATGTATGCAATGCGACAGACCTGTGATGATCCGGGCACCTCCAAATGGCCAG	573
Db	805 AAGCCCTTCTGTATGTATGCAATGCGACAGACCTGTGATGATCCGGGCACCTCCAAATGGCCAG	864
Query	574 CTTTATTACAGATTTGATCAGCTTCCATGATCAACATGTCTATCTTTCAGATC	633
Db	865 CTTTATTACAGATTTGATCAGCTTCCATGATCAACATGTCTATCTTTCAGATC	924
Query	634 AACCAACAAAACGGAGCCATTTCTCTTAACCGGAGGGATTTCAAGAAATTGAATCTGCT	693
Db	925 AACCAACAAAACGGAGCCATTTCTCTTAACCGGAGGGATTTCAAGAAATTGAATCTGCT	984
Query	694 AAGAAATCCTTCCCTTAATCTGTGATCTCAGTGAAGGACATGGAGAGCCAGAGTGAAT	753
Db	985 AAGAAATCCTTCCCTTAATCTGTGATCTCAGTGAAGGACATGGAGAGCCAGAGTGAAT	1044
Query	754 TCCTTACGTATACACATCTGTGATATCATATGACAGAGAAATTTTGGAAAGCACCA	813
Db	1045 TCCTTACGTATACACATCTGTGATATCATATGACAGAGAAATTTTGGAAAGCACCA	1104
Query	814 AAACCTGTGAGATGATGATGAGAAAATCTCAACGATCCCTCAACAAATTAATCTCAGGTG	873
Db	1105 AAACCTGTGAGATGATGATGAGAAAATCTCAACGATCCCTCAACAAATTAATCTCAGGTG	1164
Query	874 CGGTGGAATGATCCCGGTGACATATTTCTTATGTTGACAAAGAGAGCTCCAAAGATTC	933
Db	1165 CGGTGGAATGATCCCGGTGACATATTTCTTATGTTGACAAAGAGAGCTCCAAAGATTC	1224
Query	934 CCATTTTCAAATGACACAGAAAGAGATATTTACGTGACTCAGCCCTTGGACCGAGAAAGAA	993
Db	1225 CCATTTTCAAATGACACAGAAAGAGATATTTACGTGACTCAGCCCTTGGACCGAGAAAGAA	1284
Query	994 AAGGATCAATATGTTTTTATGACGTTGCAAAAGATAGATACGGAAGAAACCACTTCAAT	1053
Db	1285 AAGGATCAATATGTTTTTATGACGTTGCAAAAGATAGATACGGAAGAAACCACTTCAAT	1344

QY 1054 CCGCTGGAATTCATGTAAGTAAAGTAAATATTAATATCACTACAGTCGCTCA 1113
DB 1345 CCGCTGGAATTCATGTAAGTAAAGTAAATATTAATATCACTACAGTCGCTCA 1404
QY 1114 CCAAGTAAACGATTTTGGAGTCCAGAGATGAAGCAAGTATCGGGACCTT 1173
DB 1405 CCAAGTAAACGATTTTGGAGTCCAGAGATGAAGCAAGTATCGGGACCTT 1464
QY 1174 ACTGCATGACAGAGGATGAAGAAATATCTGCCAAGTATTCCTAATCTAAGATTTGTG 1233
DB 1465 ACTGCATGACAGAGGATGAAGAAATATCTGCCAAGTATTCCTAATCTAAGATTTGTG 1524
QY 1234 GAGCAAACTCCCAACTCCCAAGTATGAGTCTCTTCTAATCCAACTTAAGTGAAGT 1293
DB 1525 GAGCAAACTCCCAACTCCCAAGTATGAGTCTCTTCTAATCCAACTTAAGTGAAGT 1584
QY 1294 TTAAGTAACTAAACAGTCTCTTGAAGAGCAAGTATCTCTGATCAACCTTAAGCAATA 1353
DB 1585 TTAAGTAACTAAACAGTCTCTTGAAGAGCAAGTATCTCTGATCAACCTTAAGCAATA 1644
QY 1354 GAGGTGTCTGACAAAGATTTTCAAGACCTTTTGTGCAATCAACCTTAAGTATTC 1413
DB 1645 GAGGTGTCTGACAAAGATTTTCAAGACCTTTTGTGCAATCAACCTTAAGTATTC 1704
QY 1414 AATGATGAGATCCCACTCTTGAAGAAATCAAGTATTAAGAAACCTTAAGTGAAGT 1473
DB 1705 AATGATGAGATCCCACTCTTGAAGAAATCAAGTATTAAGAAACCTTAAGTGAAGT 1764
QY 1474 ACAAACTTGGGATCCCACTCTTGAAGAAATCAAGTATTAAGAAACCTTAAGTGAAGT 1533
DB 1765 ACAAACTTGGGATCCCACTCTTGAAGAAATCAAGTATTAAGAAACCTTAAGTGAAGT 1824
QY 1534 GGGAGTCTAAATTTCTGATCATATCAATTAAGGGAAGTGAAGGAGCGCTGGGGGTT 1593
DB 1825 GGGAGTCTAAATTTCTGATCATATCAATTAAGGGAAGTGAAGGAGCGCTGGGGGTT 1884
QY 1594 GAGCAGATCCCACTTACCAACCGGATATGCTAATTAAGAAACCTTAAGTGAAGT 1653
DB 1885 GAGCAGATCCCACTTACCAACCGGATATGCTAATTAAGAAACCTTAAGTGAAGT 1944
QY 1654 ACAAGCAGTCTTCCCACTTGAAGTGAAGCAAGAAATCTGAGCTTGAAGTGGT 1713
DB 1945 ACAAGCAGTCTTCCCACTTGAAGTGAAGCAAGAAATCTGAGCTTGAAGTGGT 2004
QY 1714 GTGAAGTACAAATGCAAGTCTTTTGGCAAGTTCACGCTTATTTGACAGATGAAGTAA 1773
DB 2005 GTGAAGTACAAATGCAAGTCTTTTGGCAAGTTCACGCTTATTTGACAGATGAAGTAA 2064
QY 1774 GCACCTCAATTTTCCCAACGATATCCAAAGCAAGTGAAGTGAAGTGAAGTGAAGT 1833
DB 2065 GCACCTCAATTTTCCCAACGATATCCAAAGCAAGTGAAGTGAAGTGAAGTGAAGT 2124
QY 1834 ACTAAAGTGGGCAATGTAAGTCCCAAGGATCCAGAAAGTCTGACATTAAGCTATTCAGT 1893
DB 2125 ACTAAAGTGGGCAATGTAAGTCCCAAGGATCCAGAAAGTCTGACATTAAGCTATTCAGT 2184
QY 1894 AGGGGAGACACAAAGTGGCTTAAATTAAGCAAGTGAAGTGAAGTGAAGTGAAGT 1953
DB 2185 AGGGGAGACACAAAGTGGCTTAAATTAAGCAAGTGAAGTGAAGTGAAGTGAAGT 2244
QY 1954 GCTCATTGACACAAAGGAGGAGTCCATATGAGGAGTGAAGTGAAGTGAAGTGAAGT 2013
DB 2245 GCTCATTGACACAAAGGAGGAGTCCATATGAGGAGTGAAGTGAAGTGAAGTGAAGT 2304
QY 2014 GGGGGGCTTCTTCTTAAAGTCTGTGACAGATTTCAACCTGATCTTAAGTGAAGTGAAGT 2073
DB 2305 GGGGGGCTTCTTCTTAAAGTCTGTGACAGATTTCAACCTGATCTTAAGTGAAGTGAAGT 2364
QY 2074 AACCTTCCAGAGTGAAGCAAGTCAAGGAGTGTGCTTCTTCTGCAATCCCTCAAGTGA 2133
DB 2365 AACCTTCCAGAGTGAAGCAAGTCAAGGAGTGTGCTTCTTCTGCAATCCCTCAAGTGA 2424
QY 2134 CCTGGAAGTCTCATTTTCTGAGGCTAAGTATGATGATGACACTTAATTTCTGGGGTCCCAT 2193

DB 2425 CCTGGAAGTCTCATTTTCTGAGGCTAAGTATGATGATGACACTTAATTTCTGGGGTCCCAT 2484
QY 2194 TTTACATTTTCCCTCGGCACTGAGAGCTTAACAAACGACTGGGAATTTCCAAATCAAT 2253
DB 2485 TTTACATTTTCCCTCGGCACTGAGAGCTTAACAAACGACTGGGAATTTCCAAATCAAT 2544
QY 2254 GGTACTCATGCCGACGTGCTACAGGACACAGACTTTGAGAGAGGAGGAGTGTGCTC 2313
DB 2545 GGTACTCATGCCGACGTGCTACAGGACACAGACTTTGAGAGAGGAGGAGTGTGCTC 2604
QY 2314 TTGATCCGATCAATGATGAGGAGTGGCCACCTTGAAGAGCAATTTCTTTACAGTT 2373
DB 2605 TTGATCCGATCAATGATGAGGAGTGGCCACCTTGAAGAGCAATTTCTTTACAGTT 2664
QY 2374 ACATTTCTGACATTTGTGAGAGGAATTTTCCGGCCAGAGCTGACAGACTGGAGTA 2433
DB 2665 ACATTTCTGACATTTGTGAGAGGAATTTTCCGGCCAGAGCTGACAGACTGGAGTA 2724
QY 2434 CCCACTGTGGGACATGGCAGTGTGATGCTGACACCTTCTGATGATGATTAAT 2493
DB 2725 CCCACTGTGGGACATGGCAGTGTGATGCTGACACCTTCTGATGATGATTAAT 2784
QY 2494 TTGACAGTTGTGTTTATCCGCAATTAAGAGATTAAGGCAAGATATGTTGAAGTGT 2553
DB 2785 TTGACAGTTGTGTTTATCCGCAATTAAGAGATTAAGGCAAGATATGTTGAAGTGT 2844
QY 2554 CAAGCATCTGAAGTCAACCTCTGAGAGACTGA 2586
DB 2845 CAAGCATCTGAAGTCAACCTCTGAGAGACTGA 2877

RESULT 10

US-09-833-263-1086
; Sequence 1086, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1086
; LENGTH: 2877
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(2877)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-263-1086

Query Match 72.7%; Score 2432.6; DB 10; Length 2877;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2432; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 154 CAAGAGGGGAGGATTTAGTGAACCCCTGAAGCCCATGACATTTCTATTATGAAGGCCAA 213
DB 445 CAAGAGGGGAGGATTTAGTGAACCCCTGAAGCCCATGACATTTCTATTATGAAGGCCAA 504
QY 214 GAACCGAGTCAATTTATTTCCAGTTTAAGGCAATCTCTGCTGATCTTTGAAGTGA 273
DB 505 GAACCGAGTCAATTTATTTCCAGTTTAAGGCAATCTCTGCTGATCTTTGAAGTGA 564
QY 274 ACTGGGAGACAGACAAATATTTGTATGAAGCGGAGGAGCTTCTGATTAACAACAGA 333
DB 565 ACTGGGAGACAGACAAATATTTGTATGAAGCGGAGGAGCTTCTGATTAACAACAGA 624

QY 334 GCCTTGGACAGGGAACAAGATCTACTACAAATCTCCAGGTTGGAGCCCTGGACGCTAAT 393
DB 625 GCCTTGGACAGGGAACAAGATCTACTACAAATCTCCAGGTTGGAGCCCTGGAGCTAAT 684
QY 394 GGAATTTATAGTGAAGGGTCCAGTCCCTATCAACATAGAAAGTGAAGCATCAACGACAAAT 453
DB 685 GGAATTTATAGTGAAGGGTCCAGTCCCTATCAACATAGAAAGTGAAGCATCAACGACAAAT 744
QY 454 CGACCCAGTTCCTCCAGTCAAGATAGAAAGGCTCAATTAAGGCGAACTCTCCGCCGAGA 513
DB 745 CGACCCAGTTCCTCCAGTCAAGATAGAAAGGCTCAATTAAGGCGAACTCTCCGCCGAGA 804
QY 514 AAGCCCTTCTTGTATGTCAATGCGACAGACCTGGATATCGGCGCACTCCCAATGGCGAG 573
DB 805 AAGCCCTTCTTGTATGTCAATGCGACAGACCTGGATATCGGCGCACTCCCAATGGCGAG 864
QY 574 CTTTATTTACAGATGTTCATCTCAGCTTCCATGATCAACATGTCAATGTCATTTAGATC 633
DB 865 CTTTATTTACAGATGTTCATCTCAGCTTCCATGATCAACATGTATGTATCTTTCAGATC 924
QY 634 AACACAAAAGGAGGCAATCTCTTAACCCGAGAGGATCTCGAGAAATTTGAATCCCTGCT 693
DB 925 AACACAAAAGGAGGCAATCTCTTAACCCGAGAGGATCTCGAGAAATTTGAATCCCTGCT 984
QY 694 AAGAAATCCTTCTTATATCTGTGATCTCAGTGAAGGACATGGAGGCGCAGAGTGAAT 753
DB 985 AAGAAATCCTTCTTATATCTGTGATCTCAGTGAAGGACATGGAGGCGCAGAGTGAAT 1044
QY 754 TCCTTCAGTATACCAATCTGTGATATCAATAGTGAAGAGATATTTGGAAGACCA 813
DB 1045 TCCTTCAGTATACCAATCTGTGATATCAATAGTGAAGAGATATTTGGAAGACCA 1104
QY 814 AAACCTGTGAGATGTGGAAGAACTCAACATGCTTCAACCCCATCAAAATCACTCAGGTG 873
DB 1105 AAACCTGTGAGATGTGGAAGAACTCAACATGCTTCAACCCCATCAAAATCACTCAGGTG 1164
QY 874 CGGTGGAATATCCCGGTGCAAAATATTCCTTAGTTGACAAAGAGAGTCCCAAGATTC 933
DB 1165 CGGTGGAATATCCCGGTGCAAAATATTCCTTAGTTGACAAAGAGAGTCCCAAGATTC 1224
QY 934 CCAATTTCAATTTGACAGAGAAAGATATTTACGTGACCTGAGCCGAGAGAA 993
DB 1225 CCAATTTCAATTTGACAGAGAAAGATATTTACGTGACCTGAGCCGAGAGAA 1284
QY 994 AAGGATGCAATATTTTATGAGTTGCAAGATGAGTGAAGGAAACCACTTCAAT 1053
DB 1285 AAGGATGCAATATTTTATGAGTTGCAAGATGAGTGAAGGAAACCACTTCAAT 1344
QY 1054 CCGCTGGAATTTGATGTAAGATTTAAAGATATTATGATATCACTACATGCTCGTCA 1113
DB 1345 CCGCTGGAATTTGATGTAAGATTTAAAGATTTATGATATCACTACATGCTCGTCA 1404
QY 1114 CCAATTTCAATTTGAGTTCAGAGAGATGAACGACTGGGTAAACGATATCGGAGCCCTT 1173
DB 1405 CCAATTTCAATTTGAGTTCAGAGAGATGAACGACTGGGTAAACGATATCGGAGCCCTT 1464
QY 1174 ACTGCAATGACAGGATGGAAGAAATACCTGCCAAGTTTCTTAACTACAGATTTGTG 1233
DB 1465 ACTGCAATGACAGGATGGAAGAAATACCTGCCAAGTTTCTTAACTACAGATTTGTG 1524
QY 1234 GAGCAAACTCCAAACTTCCATGATGAGTCTTCCATATCCAAACCTATGCTGGAGATG 1293
DB 1525 GAGCAAACTCCAAACTTCCATGATGAGTCTTCCATATCCAAACCTATGCTGGAGATG 1584
QY 1294 TTACAGTTAGCTTAAACAGTCTTTGAAGAGCAAGATATCTCTCAGTACAACTTAAAGATA 1353
DB 1585 TTACAGTTAGCTTAAACAGTCTTTGAAGAGCAAGATATCTCTCAGTACAACTTAAAGATA 1644
QY 1354 GAGGTGTGACAAAGATTTCAAGACCTTTGTTTGTGCAATCAACGTTATTTGATATC 1413
DB 1645 GAGGTGTGACAAAGATTTCAAGACCTTTGTTTGTGCAATCAACGTTATTTGATATC 1704

QY 1414 AATGATCAGATCCCATCTTTGAAAAATCAGATTATGAAACCTGACTCTTGTGAGAC 1473
DB 1705 AATGATCAGATCCCATCTTTGAAAAATCAGATTATGAAACCTGACTCTTGTGAGAC 1764
QY 1474 ACAACATTTGGGTCCACATCTTAAACATCCAGGCCATGATGCTGATGAGCCATTTACT 1533
DB 1765 ACAACATTTGGGTCCACATCTTAAACATCCAGGCCATGATGCTGATGAGCCATTTACT 1824
QY 1534 GGGAGTTCTAAATTTCTGATATCAATCAATTAAGGAGACAGTGAAGGAGCCCTGGGGGTT 1593
DB 1825 GGGAGTTCTAAATTTCTGATATCAATCAATTAAGGAGACAGTGAAGGAGCCCTGGGGGTT 1884
QY 1594 GACACAGATCCCATCAACACCGGATATGTATATTTAAAGGCTCTTGAATTTGAA 1653
DB 1885 GACACAGATCCCATCAACACCGGATATGTATATTTAAAGGCTCTTGAATTTGAA 1944
QY 1654 ACAGCAGCTGTTCCAAATGTGTTCOAAGAGAAATCTGAGCCCTGATGTTGGT 1713
DB 1945 ACAGCAGCTGTTCCAAATGTGTTCOAAGAGAAATCTGAGCCCTGATGTTGGT 2004
QY 1714 GTGAATTAACATGCAATGTTCTTTGCAAGTTCAAGCTTATGTCACATGTAATGAA 1773
DB 2005 GTGAATTAACATGCAATGTTCTTTGCAAGTTCAAGCTTATGTCACATGTAATGAA 2064
QY 1774 GCACCTCAATTTTCCCAACAGTATTCGAAGGCAAGTCAAGTGAAGTATGATAGGC 1833
DB 2065 GCACCTCAATTTTCCCAACAGTATTCGAAGGCAAGTCAAGTGAAGTATGATAGGC 2124
QY 1834 ACTAAAGTGGCAATGTAATGTCGAAGATTCGAAGAGTCTGGAATTAAGTATTCAGT 1893
DB 2125 ACTAAAGTGGCAATGTAATGTCGAAGATTCGAAGAGTCTGGAATTAAGTATTCAGT 2184
QY 1894 AGGGGAGACAAAGAGTGGCTTAAATTTGACACAGTGAATGCTGATGATGATG 1953
DB 2185 AGGGGAGACAAAGAGTGGCTTAAATTTGACACAGTGAATGCTGATGATGATG 2244
QY 1954 GCTCCATTTGACAGAGAACCCGAGTCCATATCGGGTCAAGTGTGAGCAGAGAAATG 2013
DB 2245 GCTCCATTTGACAGAGAACCCGAGTCCATATCGGGTCAAGTGTGAGCAGAGAAATG 2304
QY 2014 GGGGGGTCTTCTTAAAGTCTGTGTGACAGTTTCACTGATCTTATGATGATGATGAC 2073
DB 2305 GGGGGGTCTTCTTAAAGTCTGTGTGACAGTTTCACTGATCTTATGATGATGATGAC 2364
QY 2074 AACCTCCAGAGTACAGGACATCAAGGCTGTTCTTCCATCCCTCACTGATG 2133
DB 2365 AACCTCCAGAGTACAGGACATCAAGGCTGTTCTTCCATCCCTCACTGATG 2424
QY 2134 CCTGAAAGTCTCAATTTGAGGCTACTGATGATGATCAGACCTTATTTGAGGGTCCCAT 2193
DB 2425 CCTGAAAGTCTCAATTTGAGGCTACTGATGATGATCAGACCTTATTTGAGGGTCCCAT 2484
QY 2194 TTTACATTTTCCCTGAGCAGTGAAGCTTACAAAAGCACTGGGAAGTTTCAAAATCAAT 2253
DB 2485 TTTACATTTTCCCTGAGCAGTGAAGCTTACAAAAGCACTGGGAAGTTTCAAAATCAAT 2544
QY 2254 GGTATCTATGCCCCGACTGTCTACAGGCAACAGACTTTGAGAGAGGGCGTATGTCTC 2313
DB 2545 GGTATCTATGCCCCGACTGTCTACAGGCAACAGACTTTGAGAGAGGGCGTATGTCTC 2604
QY 2314 TTGATCCGATCAATGATGAGGGGTGCGGCCACCTTGAAGGATGTTCTTACACAGT 2373
DB 2605 TTGATCCGATCAATGATGAGGGGTGCGGCCACCTTGAAGGATGTTCTTACACAGT 2664
QY 2374 ACATCTGAGTGTGTGAGAGAAAGTTTTCGCGCAGCAGTCAACGACTGGAGATA 2433
DB 2665 ACATCTGAGTGTGTGAGAGAAAGTTTTCGCGCAGCAGTCAACGACTGGAGATA 2724
QY 2434 CCCACTGTGGGCAATGGCACTGTGATATCTGTCGACCACTTCTGCTGATGTTGATAT 2493
DB 2725 CCCACTGTGGGCAATGGCACTGTGATATCTGTCGACCACTTCTGCTGATGTTGATAT 2784
QY 2494 TTAGCAGTTGTGTTATCCGATTAAGAGATTAAGGCAAGATATGTTGAAGTGTCT 2553

CURRENT APPLICATION NUMBER: US/09/878.178
CURRENT FILING DATE: 2001-06-08
NUMBER OF SEQ ID NOS: 2237
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1019
LENGTH: 520
TYPE: DNA
ORGANISM: Homo sapiens
US-09-878-178-1019

Query Match 15.4%; Score 514.2; DB 9; Length 520;
Best Local Similarity 99.4%; Pred. No. 1,7e-133;
Matches 516; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

2257 ACTCATGCCGACCTGTCTACAGGACACAGACTTTGAGGAGGGGATATGCTCTTG 2316
1 ACTCATGCCGACCTGTCTACAGGACACAGACTTTGAGGAGGGGATATGCTCTTG 60
2317 ATCCGATCAATATGAGGGGTCGCGCACCCCTTGGAGGCAATTTCTTTACAGTTACA 2376
61 ATCCGATCAATATGAGGGGTCGCGCACCCCTTGGAGGCAATTTCTTTACAGTTACA 120
2377 TTCTGCAGTTGTGTGGAAGAAATTTTTCGCGCAGAGGTCAACAGCTGGGATACC 2436
121 TTCTGCAGTTGTGTGGAAGAAATTTTTCGCGCAGAGGTCAACAGCTGGGATACC 180
2437 ACTGTGGGCAATGAGGAGTTGATATCTGCTGACCACTTGTGATTTGATATTTA 2496
181 ACTGTGGGCAATGAGGAGTTGATATCTGCTGACCACTTGTGATTTGATATTTA 240
2497 GCAGTTGTGTTATCCGCAATAAAGAGATTAAGGCAAGATTAATTTGAAAGTCTCAA 2556
241 GCAGTTGTGTTATCCGCAATAAAGAGATTAAGGCAAGATTAATTTGAAAGTCTCAA 300
2557 GCATCTGAAGTCAAACTCTGAGAGAGTGAATTTGAAAAGATTTGATATTTAG 2616
301 GCATCTGAAGTCAAACTCTGAGAGAGTGAATTTGAAAAGATTTGATATTTAG 360
2617 CAAGTGTATTTAGCAACAACCATCTCATCTTATTAATCTTAATCAAGTGCATATA 2676
361 CAAGTGTATTTAGCAACAACCATCTCATCTTATTAATCTTAATCAAGTGCATATA 420
2677 ATTTTAAACAGATATTCCTCTTGTCTTAAATATTTCTTTTGTAG 2736
421 ATTTTAAACAGATATTCCTCTTGTCTTAAATATTTCTTTTGTAG 480
2737 GTGAGTCTTGTCTGTGCGCCAGGCTGAGATACAGTGG 2775
481 GTGAGTCTTGTCTGTGCGCCAGGCTGAGATACAGTGG 519

RESULT 14
US-10-146-502-1019
Sequence 1019, Application US/10146502
Publication No. US2003069180A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Harlocker, Susan L.
APPLICANT: Secretist, Heather
APPLICANT: Wang, Ajun
APPLICANT: Stolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.527C2
CURRENT APPLICATION NUMBER: US/10/146.502
CURRENT FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 2241
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1019
LENGTH: 520
TYPE: DNA
ORGANISM: Homo sapiens
US-10-146-502-1019

Query Match 15.4%; Score 514.2; DB 9; Length 520;
Best Local Similarity 99.4%; Pred. No. 1,7e-133;
Matches 516; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

2257 ACTCATGCCGACCTGTCTACAGGACACAGACTTTGAGGAGGGGATATGCTCTTG 2316
1 ACTCATGCCGACCTGTCTACAGGACACAGACTTTGAGGAGGGGATATGCTCTTG 60
2317 ATCCGATCAATATGAGGGGTCGCGCACCCCTTGGAGGCAATTTCTTTACAGTTACA 2376
61 ATCCGATCAATATGAGGGGTCGCGCACCCCTTGGAGGCAATTTCTTTACAGTTACA 120
2377 TTCTGCAGTTGTGTGGAAGAAATTTTTCGCGCAGAGGTCAACAGCTGGGATACC 2436
121 TTCTGCAGTTGTGTGGAAGAAATTTTTCGCGCAGAGGTCAACAGCTGGGATACC 180
2437 ACTGTGGGCAATGAGGAGTTGATATCTGCTGACCACTTGTGATTTGATATTTA 2496
181 ACTGTGGGCAATGAGGAGTTGATATCTGCTGACCACTTGTGATTTGATATTTA 240
2497 GCAGTTGTGTTATCCGCAATAAAGAGATTAAGGCAAGATTAATTTGAAAGTCTCAA 2556
241 GCAGTTGTGTTATCCGCAATAAAGAGATTAAGGCAAGATTAATTTGAAAGTCTCAA 300
2557 GCATCTGAAGTCAAACTCTGAGAGAGTGAATTTGAAAAGATTTGATATTTAG 2616
301 GCATCTGAAGTCAAACTCTGAGAGAGTGAATTTGAAAAGATTTGATATTTAG 360
2617 CAAGTGTATTTAGCAACAACCATCTCATCTTATTAATCTTAATCAAGTGCATATA 2676
361 CAAGTGTATTTAGCAACAACCATCTCATCTTATTAATCTTAATCAAGTGCATATA 420
2677 ATTTTAAACAGATATTCCTCTTGTCTTAAATATTTCTTTTGTAG 2736
421 ATTTTAAACAGATATTCCTCTTGTCTTAAATATTTCTTTTGTAG 480
2737 GTGAGTCTTGTCTGTGCGCCAGGCTGAGATACAGTGG 2775
481 GTGAGTCTTGTCTGTGCGCCAGGCTGAGATACAGTGG 519

RESULT 15
US-10-066-543-3247
Sequence 3247, Application US/10066543
Publication No. US2003087818A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Pyke, Ruth A.
APPLICANT: Xu, Jiangchun
APPLICANT: Indira, Carol Yoseph
APPLICANT: Lodes, Michael J.
APPLICANT: Secretist, Heather
APPLICANT: Carter, Darick
APPLICANT: Fanger, Gary R.
APPLICANT: Smith, Carole L.
APPLICANT: Durham, Margarita
APPLICANT: Stolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.563
CURRENT APPLICATION NUMBER: US/10/066.543
CURRENT FILING DATE: 2002-01-31
NUMBER OF SEQ ID NOS: 3417
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3247
LENGTH: 517
TYPE: DNA
ORGANISM: Homo sapiens
US-10-066-543-3247

Query Match 15.3%; Score 512.4; DB 9; Length 517;
Best Local Similarity 99.8%; Pred. No. 5.4e-133;

Matches 513; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 2257 ACTATATCCCGACCTGTCTTACAGGACACAGACTTTGAGAGAGAGGCGGTATGTCTCTG 2316
    |||
Db 1 ACTCATGCCGACGTCTTACAGGACACAGACTTTGAGAGAGAGGCGGTATGTCTCTG 60
    |||
QY 2317 ATCCGATCAATGATGGGGGTGGGCAACCTTGGAGGCAATGTTCTTTACAGTTACA 2376
    |||
Db 61 ATCCGATCAATGATGGGGGTGGGCAACCTTGGAGGCAATGTTCTTTACAGTTACA 120
    |||
QY 2377 TTCTGCAATTGTGGAGAGAAATTGTTCCGSCCAGAGGTCAACAGACTGGGATACC 2436
    |||
Db 121 TTCTGCAATTGTGGAGAGAAATTGTTCCGSCCAGAGGTCAACAGACTGGGATACC 180
    |||
QY 2437 ACTGTGGGCAATGGAGTTGGTATCTGTGACCAACCTTCTGTGATTTGGTATATTTA 2496
    |||
Db 181 ACTGTGGGCAATGGAGTTGGTATCTGTGACCAACCTTCTGTGATTTGGTATATTTA 240
    |||
QY 2497 GCAGTTGTGTTATCCGCAATTAAGAGGATTAAGGCAAAAGATATGTTGAAGTCTCAA 2556
    |||
Db .241 GCAGTTGTGTTATCCGCAATTAAGAGGATTAAGGCAAAAGATATGTTGAAGTCTCAA 300
    |||
QY 2557 GCATCTGAAGTCAAACTCTGAGAAGCTGAATTGAAAAGGAAATGTTGAATTTATATAG 2616
    |||
Db 301 GCATCTGAAGTCAAACTCTGAGAAGCTGAATTGAAAAGGAAATGTTGAATTTATATAG 360
    |||
QY 2617 CAAGTGTATTTGCAACAACCATTCATTCCTTCACTTTCATCTAACGTGATATA 2676
    |||
Db 361 CAAGTGTATTTGCAACAACCATTCATTCCTTCACTTTCATCTAACGTGATATA 420
    |||
QY 2677 ATTTTAAACAGATATTCCTCTGTGCTTAAATATTTGCTAAATATTTCTTTTGGAG 2736
    |||
Db 421 ATTTTAAACAGATATTCCTCTGTGCTTAAATATTTGCTAAATATTTCTTTTGGAG 480
    |||
QY 2737 GTGAGTCTTGCTCTGTGCGCCAGGCTGAGTAC 2770
    |||
Db 481 GTGAGTCTTGCTCTGTGCGCCAGGCTGAGTAC 514
    |||
```

Search completed: June 22, 2003, 01:03:24
Job time : 453 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2003, 20:59:00 ; Search time 4130 Seconds
(without alignments)
13117.168 Million cell updates/sec

Title: US-10-025-380-1076

Perfect score: 3345

Sequence: 1 gaattccgctcgcacacg.....cctttcgcgccgcgaattc 3345

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estbda:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	719.4	21.5	1125	14	BM918123 AGENCOURT
2	609.8	18.2	613	10	AM859603 MRI-CT035
3	594.8	17.8	625	12	BE696136 MRI-CT025
4	540.8	16.2	573	9	AI347909 qp0e04.x
5	524.8	15.7	537	9	AI732453 z190f07.x
6	516.8	15.4	540	13	BG978653 PML-CN009

7	513.6	15.4	590	10	AM859530 MRI-CT035
8	500.4	15.0	519	10	AM859550 MRI-CT035
9	500.2	15.0	519	10	AM859550 MRI-CT035
10	471.2	14.1	563	10	AM862312 RC4-CT036
11	470.4	14.1	475	9	AI262603 qk41h05.x
12	468.4	14.0	530	9	AI102326
13	467.2	14.0	472	14	BM799332 K-EST0083
14	464.2	13.9	486	13	BG981361 CM4-CN008
15	452.8	13.5	477	12	BF334481 MRI-CT035
16	451.6	13.5	495	10	AM853021
17	450.6	13.5	554	9	AA088861 z190f07.8
18	449.4	13.4	457	9	AI920859 w015b06.x
19	435.2	13.0	706	10	AM75947 RC0-CT020
20	435	13.0	455	12	BE696126 MRI-CT025
21	434.6	13.0	873	12	BF582721 602094137
22	430.8	12.9	620	10	AM77543 OVO-CT022
23	418.4	12.5	507	10	AM753130 RC1-CT024
24	410.4	12.3	464	10	AM391746 CM2-ST022
25	408.4	12.2	428	9	AI623483 t619c10.x
26	404.2	12.1	436	13	BG981357 CM4-CN008
27	401.8	12.0	406	10	AM51846 RC1-CT019
28	401.8	12.0	408	9	AI922856 w014b06.x
29	401.6	12.0	551	10	AM887172 RC1-OT008
30	399.4	11.9	461	10	AM859489 MRI-CT035
31	394.6	11.8	657	12	BF584025 602096265
32	394.4	11.8	405	13	BG980846 CM4-CN008
33	391	11.7	403	9	AA053188 z172a06.x
34	387.4	11.6	551	12	BE808624 z13876.MA
35	383.8	11.5	441	14	BQ320693 RC4-CT084
36	383.8	11.5	771	9	AI566493 t160e08.x
37	383.2	11.5	753	12	BF577684 602092567
38	377.8	11.3	435	13	BG981295 CM4-CN008
39	375.4	11.2	417	10	AM351659 OV2-CT014
40	372.4	11.1	408	14	BO320866 RC4-CT082
41	369	11.0	408	12	BF754330 IL5-CT051
42	368.6	11.0	477	10	AM859516 MRI-CT035
43	364.4	10.9	395	10	AM578054 RC3-CN001
44	364.2	10.9	447	10	AM859567 MRI-CT035
45	360.4	10.8	454	9	AA305406 EST176556

ALIGNMENTS

RESULT 1
LOCUS BM918123 1125 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6611470 NIH_MGC_106 Homo sapiens CDNA clone IMAGE:5485497
ACCESSION BM918123
VERSION BM918123.1 GI:19368502
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1125)
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINT at:
http://image.llnl.gov
Plate: LINC2016 row: h column: 10
High quality sequence stop: 662.
Location/Qualifiers
1. 1125


```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="IMAGE:5485497"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: blood; Vector: pORF1; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."
BASE COUNT      259 a      321 c      258 g      286 t      1 others
ORIGIN
Query Match      21.5%; Score 719.4; DB 14; Length 1125;
Best Local Similarity 97.0%; Pred. No. 2.6e-128;
Matches 754; Conservative 0; Mismatches 21; Indels 2; Gaps 2;
QY 1825 GCTATAGGCACTAAGTGGGCAATGTGACTGCCAAGATCCAGAAGCTTGGACATAAGC 1884
DB 1 GCTATAGGCACTAAGTGGGCAATGTGACTGCCAAGATCCAGAAGCTTGGACATAAGC 60
QY 1885 TATTCACCTGAGGGAGACACAAAGGTTGGCTTAAATGACCACTGACTGTGAGATC 1944
DB 61 TATTCACCTGAGGGAGACACAAAGGTTGGCTTAAATGACCACTGACTGTGAGATC 120
QY 1945 TTTAGTGGCTCATTTGAGACAGAGAACCCGGAAGTCCATTCGGGTACAAAGTGGTGGC 2004
DB 121 TTTAGTGGCTCATTTGAGACAGAGAACCCGGAAGTCCATTCGGGTACAAAGTGGTGGC 180
QY 2005 ACAGAGTAGGGGGGCTCTTCTTAAGCTGTGTGCAAGTTCACTGATCTTATGAT 2064
DB 181 ACAGAGTAGGGGGGCTCTTCTTAAGCTGTGTGCAAGTTCACTGATCTTATGAT 240
QY 2065 GTGAATGACAACTCCCAAGCTAGCCAGACATCACAGGGCTTTCTTGGCATCC 2124
DB 241 GTGAATGACAACTCCCAAGCTAGCCAGACATCACAGGGCTTTCTTGGCATCC 300
QY 2125 CTCAGTGAACCTGGAAGTCTCATTTTGGAGCTACTGATGATGATGATGATGATGAT 2184
DB 301 CTCAGTGAACCTGGAAGTCTCATTTTGGAGCTACTGATGATGATGATGATGATGAT 360
QY 2185 GGTCCCATTTTACATTTTCCCTCGGAGTGAAGCTTAAACAAAGCACTGGAAAGTTCC 2244
DB 361 GGTCCCATTTTACATTTTCCCTCGGAGTGAAGCTTAAACAAAGCACTGGAAAGTTCC 420
QY 2245 AAAATCAATGATCTACTGCCCCGACTGTCTACAGGACACAGACTTTGAGAGAGGGCG 2304
DB 421 AAAATCAATGATCTACTGCCCCGACTGTCTACAGGACACAGACTTTGAGAGAGGGCG 480
QY 2305 TATGTCGCTTGATCCGCATCAATGATGGGGGTGGCCACCTTGAAGGATTTCT 2364
DB 481 TATGTCGCTTGATCCGCATCAATGATGGGGGTGGCCACCTTGAAGGATTTCT 540
QY 2365 TTACCACTTACATTTCTGCAAGTTGTGTGAGAGAAAGTTGTTTCCGGCCACAGGTCCACAG 2424
DB 541 TTACCACTTACATTTCTGCAAGTTGTGTGAGAGAAAGTTGTTTCCGGCCACAGGTCCACAG 600
QY 2425 ACTGGGATACCCACTGTGGGAGATGGCACTGTGATACCTGACACACCTTCTGATAT 2484
DB 601 ACTGGGATACCCACTGTGGGAGATGGCACTGTGATACCTGACACACCTTCTGATAT 660
QY 2485 GGTAAATTTTAAAG-CAAGTTGTGTTATCCGCATTAAGAGAGATTAAGGATTAATGT 2543
DB 661 GGTAAATTTTAAAGCAAGTTGTGCGCTATCCGCTTAAGAGAGATTAAGGATTAATGT 720
QY 2544 TGAAGTCTCAAGCATC-TGAAGTCAAACTCTGAGAGTGAATTTGAAAGAGAA 2599
DB 721 TGAAGTCTCAAGCCTCTTGAAGCAAACTCTGAGAGTGAATTTGAAAGAGAA 777

```

```

RESULT 2
LOCUS      AW859603      613 bp      mRNA      linear      EST 19-MAY-2000
DEFINITION MW1-CT0355-180200-007-e07 CT0355 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW859603
VERSION     AW859603.1
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 613)
            Dias Neto, E., Garcia Correa, R., Verjovsky-Almeida, S., Brites, M.R.,
            Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
            Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
            Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
            , M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
            Simpson, A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
            2020263.
            Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?l=622-MW1-CT0355-180
            200-007-e07&t3=2000-02-18&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 613.
FEATURES
    source
        1..613
            location/Qualifiers
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_id="CT0355"
                /dev_stage="Adult"
                /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
                SmaI; A mini-library was made by cloning products derived
                from ORESTES PCR (U.S. Letters Patent Application No. 196
                ,716 - Ludwig Institute for Cancer Research) profiles
                into the puc 18 vector. Reverse transcription of tissue
                mRNA and cDNA amplification were performed under low
                stringency conditions."
BASE COUNT      195 a      125 c      134 g      159 t
ORIGIN
Query Match      18.2%; Score 609.8; DB 10; Length 613;
Best Local Similarity 99.7%; Pred. No. 3.4e-107;
Matches 611; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1318 AAGAGCAGATACTCTCTCACTGATCACTTACGATAGAGGTGCTGCAAAAGTTCAAG 1377
DB 1 AAGAGCAGATACTCTCTCACTGATCACTTACGATAGAGGTGCTGCAAAAGTTCAAG 60
QY 1378 ACCCTTTGTTTGTGCAAACTCAAGCTTATGATATCAATGATATCCCATCTTTGAA 1437
DB 61 ACCCTTTGTTTGTGCAAACTCAAGCTTATGATATCAATGATATCCCATCTTTGAA 120
QY 1438 AAATCAGATTAATGAAACCTGACTCTTGTGAGAGACAAACATTTGGTCCACATCTTA 1497
DB 121 AAATCAGATTAATGAAACCTGACTCTTGTGAGAGACAAACATTTGGTCCACATCTTA 180
QY 1498 ACCATCAGGCCACTGATGCTGATGAGCCATTTACTGGAGTTCTAAATTCGTATCAT 1557
DB 181 ACCATCAGGCCACTGATGCTGATGAGCCATTTACTGGAGTTCTAAATTCGTATCAT 240

```

QY	1558	ATCATTAAGGAGGACAGTACGAGGACCGCTGGGGGTTGACACAGATCCCATTACCAACAC	1617
Db	241	ATCATTAAGGAGGACAGTACGAGGACCGCTGGGGGTTGACACAGATCCCATTACCAACCC	300
QY	1618	GGATATGTCATTAATTAATAAAGCCTCTTGATTTTGAACAGCAGCGCTGTTTCCAACTGTG	1677
Db	301	GGATATGTCATTAATTAATAAAGCCTCTTGATTTTGAACAGCAGCGCTGTTTCCAACTGTG	360
QY	1678	TTCAAAGCAGAAAAATCCTGAGCTTCTAGTGTGTTGGTGTGAAGTACATAGCAAGTCTTTT	1737
Db	361	TTCAAAGCAGAAAAATCCTGAGCTTCTAGTGTGTTGGTGTGAAGTACAAAGCAAGTCTTTT	420
QY	1738	GCAAAGTTCACGCTTAATGTGACAGATGNGATGAAGCACCCTCAATTTTCCCAACAGTA	1797
Db	421	GCCAGTTTCACGCTTAATGTGACAGATGTGAATGAAGCACCCTCAATTTTCCCAACAGTA	480
QY	1798	TTCCAGCGAAAGTCAGTGAAGATGTAGCTTAGGCACTAAAGTGGGCAATAGTACCTCC	1857
Db	481	TTCCAGCGAAAGTCAGTGAAGATGTAGCTTAGGCACTAAAGTGGGCAATAGTACCTCC	540
QY	1858	AAGGATCCAGAAAGTCTGGAACATAAGCTATTCACTGAGGGGAGACACAAAGGTTGGCTT	1917
Db	541	AAGGATCCAGAAAGTCTGGAACATAAGCTATTCACTGAGGGGAGACACAAAGGTTGGCTT	600
QY	1918	AAAATTGACCAAG 1930	
Db	601	AAAATTGACCAAG 613	

RESULT 3			
LOCUS	BE696136/c.	625 bp	linear
DEFINITION	MR_1C0258-040700-002-e10	CT0258	Homo sapiens CDNA, mRNA sequence.
ACCESSION	BE696136		
VERSION	BE696136.1	GI:10083296	

ORGANISM
Homo, sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominae; Homo.

REFERENCES

1 (pages 1 to 625)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Bionces, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matukuma, A., Bata, G.S., Simpson, D.H., Brunslein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
COMMENT	Contact: Stimpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@mdwlg.org.br
This sequence was derived from the FAPSP/ICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.mdwlg.org.br/scripts/gethtml2.pl?l1=et2-MR1-CT0258-040700-002-El0&t3=2000-07-04&t4=1>)
Seq primer: puc 18 forward
High quality sequence stop: 620.

FEATURES	Location/Qualifiers
source	1. .625

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_idb="CT0258"
/dev_stage="Adult"
/ncbi="Organ: colon; Vector: puc19; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived

```

from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT	143 a	153 g	188 t
ORIGIN	141 c		

Query Match	17.8%	Score 594.8	DB 12	Length 625
Best Local Similarity	98.8%	Pred. No. 2.6e-104		
Matches 599	Conservative 0	Mismatches 7	Indels 0	Gaps 0

OY	277	GGGGAGACAGCAATATTTTGTATGAAAGGGAGGGACCTTCTGTATTAACAAGACC	336
Db	606	GGGAGACAGCAATATTTTGTATGAAAGGGAGGGACCTTCTGTATTAACAAGACC	547
OY	337	TTGGACAGGGAAACAAGATCTTACTACATCTCCAGGTTGCAGCCCTTGACGCTATGGA	396
Db	546	TTGGACAGGGAAACAAGATCTTACTACATCTCCAGGTTGCAGCCCTTGACGCTATGGA	487
OY	397	ATTATAGTGAAGGCTCCAGTCCCTATATCACTAAGATGGAAGACATCAAGACATGGA	456
Db	486	ATTATAGTGAAGGCTCCAGTCCCTATATCACTAAGATGGAAGACATCAAGACATGGA	427
OY	457	CCCAAGTTTCCAGTCAAAGTACGAAGGCTCAGTAAGGACAGAACTCTCGCCACAGAAAG	516
Db	426	CCCAAGTTTCCAGTCAAAGTACGAAGGCTCAGTAAGGACAGAACTCTCGCCACAGAAAG	367
OY	517	CCCTTCTTGTATGTCAATGACCAGACCTTGATATCGGACCACCTCCCAATGGCCAGCTT	576
Db	366	CCCTTCTTGTATGTCAATGACCAGACCTTGATATCGGACCACCTCCCAATGGCCAGCTT	307
OY	577	TATTACAGAGATGTGCATCCAGCTTCCCATGATGCAACAAATGTCAATGTAATTTCAATCAAC	636
Db	306	TATTACAGAGATGTGCATCCAGCTTCCCATGATGCAACAAATGTATGTAATTTCAATCAAC	247
OY	637	AACAAAACGGAGGCATCTCTCTTACCCGAGAGGATCTCAGAAATGAAATCTGTCTAAG	696
Db	246	AACAAAACGGAGGCATCTCTCTTACCCGAGAGGATCTCAGAAATGAAATCTGTCTAAG	187
OY	697	AATCTCTCTTAATCTGTGTGATCTCAGTGAAGACATGGAGGCCAGAGTGAAGATTC	756
Db	186	AATCTCTCTTAATCTGTGTGATCTCAGTGAAGACATGGAGGCCAGAGTGAAGATTC	127
OY	757	TTCAAGTATCAACATCTGTGATATCATATGTAACAGAGATATTTTGAAGAAGCAACAA	816
Db	126	TTCAAGTATCAACATCTGTGATATCATATGTAACAGAGATATTTTGAAGAAGCAACAA	67
OY	817	CCTGTGAGATGTGGAAGAACTCAACTGATCTCACCACCAATCAAAATCATTCAAGTGGG	876
Db	66	CCTGTGAGATGTGGAAGAACTCAACTGATCTCACCACCAATCAAAATCATTCAAGTGGG	7
OY	877	TGGAAT 882	
Db	6	TGGAAT 1	

RESULT 4	LOCUS	DEFINITION	ACCESSION	VERSION
AI347909/c	AI347909	573 bp mRNA	AI347909	AI347909.1
		NCI CGAP C98 Homo sapiens cDNA clone IMAGE:1927422 3' similar to TR:Q55356 Q15356 PI-CADHERIN. [1] '', mRNA sequence.		GI:4085115

SOURCE ORGANISM

REFERENCE
AUTHORS
TITLE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Butelostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 573)
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ccgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
1 (bases 1 to 540)	Dias Neto, E., Garcia Correa, R., Veitovskí-Almeida, S., Britones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P., Goldman, G.H., Carvalho, A.F., Matsumura, A., Bata, G.S., Simpson, D.H., Bruneirlein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags	Proc. Natl. Acad. Sci. U.S.A.	97 (7), 3491-3496 (2000)	
20202863	Contact: Simpson A.J.G.				

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel.: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL.
(<http://www.ludwig.org.br/scripts/gethcm12.pl?tbl=PMkt2-PM1-CNN098-020101-004-e02ecf3-2001-01-02&cf=1>)
Seq primer: puc 18 forward
High quality sequence start: 7
High quality sequence stop: 538.

```

FEATURES
    source
        Location/Qualifiers
            1..540
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_id="CN0098"
                /dev_stage="Adult"
                /note="Organ: colon normal; Vector: puc18; Site 1: SmaI;
                Site 2: SmaI; A mini-library was made by cloning products
                derived from ORS785 PCR (U.S. Letters Patent application
                No. 196,716 - Ludwig Institute for Cancer Research)
                profiles into the pUC 18 vector. Reverse transcription of
                tissue mRNA and cDNA amplification were performed under
                low stringency conditions."
    BASE COUNT
    136 a 136 c 127 g 140 t 1 others
    ORIGIN

```

.15.4%; Score 516.8; DB 13; Length 540,

Best Local Similarity 98.3%; Pred. No. 2.6e-89;
Matches 532; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY	1754	TTGGACAGAGTGAATGAAGCACTCAATTTTCCCAACAGATATTCCAAAGGAAAGTCA	1813
Dd	540	TTTGACAGATGTAAATGAAGCACTCAATTTTCCCAACAGATATTCCAAAGGAAAGTCA	481
QY	1814	GTGAGAGTGTAGCTATATGGCACTAAAGTGGCAATGTGACTGCCAAGGATCAGAAAGTCT	1872
Dd	480	GTGAGAGTGTATGTATATGGCCCTTAAAGTGGCAATGTGACTGCCAAGGATCAGAAAGTCT	421
QY	1874	TGACATATAGCTATTCACTGAGGGGAGACACAGAAGTTGGCTTAAATTGACCACTGA	1933
Dd	420	TGACATATAGCTATTCACTGAGGGGAGACACAGAAGTTGGCTTAAATTGACCACTGA	361
QY	1934	CTGGTGAAGATCTTATAGTGTGGCTCAATTGACAGAGAAGCCGGAAAGTCCATATGGGTAC	1993
Dd	360	CTGGTGAAGATCTTATAGTGTGGCTCAATTGACAGAGAAGCCGGAAAGTCCATATGGGTAC	301
QY	1994	AAGTGTGTGGCACAGAGTAGAGGGGGGCTCTTCTTAAGCTGTGTCTCAGAGTTCCACTGA	2053
Dd	300	AAGTGTGTGGCACAGAGTAGAGGGGGGCTCTTCTTAAGCTGTGTCTCAGAGTTCCACTGA	241
QY	2054	TCCTTATGATGTGAATGACAAACCTTCCAGGCTAGCCCAAGACTCAACGGGCTTGTCT	2113
Dd	240	TCCTTATGATGTGAATGACAAACCTTCCAGGCTAGCCCAAGACTCAACGGGCTTGTCT	181
QY	2114	TCGGCCATCCCCAGTGCACCTGGGAAGTCTCATTTTCAGAGCTATATGATATCAGC	2173
Dd	180	TCGGCCATCCCCAGTGCACCTGGGAAGTCTCATTTTCAGAGCTATATGATATATCAGC	121
QY	2174	ACTATATTCGGGGGTCCCATATTTTACATTTTCCCTCGGCAAGTGAAGCTTACAAACGACT	2233
Dd	120	ACTATATTCGGGGGTCCCATATTTTACATTTTCCCTCGGCAAGTGAAGCTTACAAACGACT	61
QY	2234	GGGAAGTTTCCAAATCAATGTGATCTATGCCCCGACTGTCTTACCAAGGCACAGACTTTG	2293
Dd	60	GGGAAGTTTCC-AAATCAATGTGATCTATGCCCCGACTGTCTTACCAATGATGATCTTG	2
QY	2294	A 2294	
Dd	1	A 1	

[illegible]

Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?l=et2=MR1-CT0355-160>)
 200-003-e05et3=2000-02-16et4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 22
 High quality sequence stop: 577.
 Location/Qualifiers

FEATURES

source

1. 590
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1lb="CT0355"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 183 a 127 c 132 g 148 t

ORIGIN

Query Match 15.4%; Score 513.6; DB 10; Length 590;
 Best Local Similarity 99.1%; Pred. No. 1e-88; 4; Indels 1; Gaps 1;
 Matches 527; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1330 ACTCCCTAGTACAACTTACAGTAGAGGTCTGACAAAGATTTCAGACCCCTTGTGTTT 1389
 DB 19 ACGCTCAGTACAACTGACAGTAGAGGTCTGACAAAG-TCAGACCCCTTGTGTTT 77

QY 1390 GTGCAATCAACGTTATTGATATCATGATATAGATCCCATCTTTGAAAAATCGATTAT 1449
 DB 78 GTGCAATCAACGTTATTGATATCATGATATAGATCCCATCTTTGAAAAATCGATTAT 137

QY 1450 GGAAACCTGACTCTTGTGTAAGACAAACATTGGGTCCACCATCTTAACCATCCAGGCC 1509
 DB 138 GGAAACCTGACTCTTGTGTAAGACAAACATTGGGTCCACCATCTTAACCATCCAGGCC 197

QY 1510 ACTGATCTGATGAGCAATTTACTGGAGTTCTTAAATTCGTATCATATCATTAAGGGA 1569
 DB 198 ACTGATCTGATGAGCAATTTACTGGAGTTCTTAAATTCGTATCATATCATTAAGGGA 257

QY 1570 GACAGTAGAGGACCGCTGGGGGTGACACAGATCCCATACCAACCGGATATGTATA 1629
 DB 258 GACAGTAGAGGACCGCTGGGGGTGACACAGATCCCATACCAACCGGATATGTATA 317

QY 1630 ATTAAAAAGCCTCTTGTGTAAGACAGACGCTGTTCCAACTTTGTTCCAAAGCAGAA 1689
 DB 318 ATTAAAAAGCCTCTTGTGTAAGACAGACGCTGTTCCAACTTTGTTCCAAAGCAGAA 377

QY 1690 AATCCTGAGCCTTAGTGTGTGTGTGTAAGTACATGCAAGTCTTTTGCCAAAGTTCAG 1749
 DB 378 AATCCTGAGCCTTAGTGTGTGTGTGTAAGTACATGCAAGTCTTTTGCCAAAGTTCAG 437

QY 1750 CTTATTGTGACAGATGTAAGTGAAGCACTCAATTTTCCCAACGTAATTCAGACGAAA 1809
 DB 438 CTTATTGTGACAGATGTAAGTGAAGCACTCAATTTTCCCAACGTAATTCAGACGAAA 497

QY 1810 GTCAAGTAGATGATGATATGAGCACTAAAGTGGGCAATGATGCAAGG 1861
 DB 498 GTCAAGTAGATGATGATATGAGCACTAAAGTGGGCAATGATGCAAGG 549

RESULT 8
 AM859550 519 bp mRNA linear EST 19-MAY-2000
 LOCUS MR1-CT0355-160200-004-b02 CT0355 Homo sapiens cDNA, mRNA sequence.
 DEFINITION
 ACCESSION AM859550
 VERSION AM859550.1 GI:7955243
 KEYWORDS
 SOURCE human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 519)

AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Coe, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.U.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

COMMENT

Contact: Simpson A.U.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?l=et2=MR1-CT0355-160>)
 200-004-b02et3=2000-02-16et4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 81
 High quality sequence stop: 517.
 Location/Qualifiers

FEATURES

source

1. 519
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1lb="CT0355"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 160 a 111 c 107 g 141 t

ORIGIN

Query Match 15.0%; Score 500.4; DB 10; Length 519;
 Best Local Similarity 97.9%; Pred. No. 3.7e-86;
 Matches 507; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1329 TACTCCAGTACAACTTAAAGTAGAGGTCTGACAAAGATTTCAGACCCCTTGTGTT 1388
 DB 2 TACTCCAGTACAACTTAAAGTAGAGGTCTGACAAAGATTTCAGACCCCTTGTGTT 61

QY 1389 TGTGCAATCAACGTTATTGATATCATGATATGATCCCATCTTTGAAAAATCGATTAT 1448
 DB 62 TGTGCAATCAACGTTATTGATATCATGATATGATCCCATCTTTGAAAAATCGATTAT 121

QY 1449 TGGAAACCTGACTCTTGTGTAAGACAAACATTGGGTCCACCATCTTAAACCATCCAGGC 1508
 DB 122 TGGAAACCTGACTCTTGTGTAAGACAAACATTGGGTCCACCATCTTAAACCATCCAGGC 181

QY 1509 CACTGATGCTGATGAGCAATTTACTGGAGTTCTTAAATTCGTATCATATCATTAAGG 1568
 DB 182 CACTGATGCTGATGAGCAATTTACTGGAGTTCTTAAATTCGTATCATATCATTAAGG 241

QY 1569 AGACAGTAGAGGACGCTGGGGGTGACACAGATCCCATACCAACCGGATATGTAT 1628
 DB 242 AGACAGTAGAGGACGCTGGGGGTGACACAGATCCCATACCAACCGGATATGTAT 301

QY 1629 AATTTAAAAAGCCTCTTGTGTAAGACAGAGCTGTTTCCAACTTGTGTTCAAAGCAGA 1688
 DB 302 AATTTAAAAAGCCTCTTGTGTAAGACAGAGCTGTTTCCAACTTGTGTTCAAAGCAGA 361

QY 1689 AATCCGAGGCTCTAGTGTGGTGTGAAGTACATGCAAGTCTTTTGGCAAGTTCAC 1748
 DB 362 AATTCGAGGCTCTAGTGTGGTGTGAAGTACATGCAAGTCTTTTGGCAAGTTCAC 421
 QY 1749 GCTATTGTGACAGATGTGAAGCAAGCCTTCATTTTCCCAACAGGATTTCCAGGAA 1808
 DB 422 GCTATTGTGACAGATGTGAAGCAAGCCTTCATTTTCCCAACAGGATTTCCAGGAA 481
 QY 1809 AGTCAGTACGATGTAGCTATAGCACTAAAGTGAGCA 1846
 DB 482 AGTCAGTACGATGTAGCTATAGCACTAAAGTGAGCA 519

RESULT 9
 A1733765 514 bp mRNA linear EST 14-JUN-1999
 LOCUS 2190f07.Y5 StrataGene colon (#937204) Homo sapiens cDNA clone
 DEFINITION IMAGE:511909.5, similar to TR:Q12864 Q12864 INTESTINAL
 PEPTIDE-ASSOCIATED TRANSPORTER HPT-1. [1] ; mRNA sequence.
 A1733765
 VERSION A1733765.1 GI:5054878
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 514)
 NCBI/NIH-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute / National Institute of Dental Research,
 Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Other ESTs: z190f07.x5
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 This clone is available royalty-free through LBNL; contact the
 IMAGE Consortium (infoimage.lbnl.gov) for further information.
 This read is a RESSEQUENCE of a previously sequenced human clone
 Original clone citation: Washu-Merck EST Project
 This read has been verified (found to hit its original self in the
 correct orientation)
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40RP from Gldco
 High quality sequence stop: 405.
 Location/Qualifiers
 1..514
 /organism="Homo sapiens"
 /db_xref="GDB:384404"
 /db_xref="taxon:9606"
 /clone="IMAGE:511909"
 /clone_1db="Stratagene colon (#937204)"
 /tissue type="tumor"
 /cell_line="T84 carcinoma cell line"
 /lab_host="SOLR cells (kanamycin resistant)"
 /note="Organ: colon; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
 Oligo dT. T-84 colonic epithelial cell line. Average
 insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
 sequence: 5' GAATTGGCAGCG 3' ~3' adaptor sequence: 5'
 CTCGAGTTTCTTTTCTTTTCTTTT 3' "

BASE COUNT 135 a 115 c 132 g 132 t
 ORIGIN

Query Match 15.0%; Score 500.2; DB 9; Length 514;
 Best Local Similarity 98.4%; Pred. No. 4e-86; Indels 0; Gaps 0;
 Matches 505; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1677 GTTCAAGCAGAAATCTGAGCCTCTAGTGTGGTGTGAAGTACATGCAAGTCTTT 1736
 DB 1 GTTCAAGCAGAAATCTGAGCCTCTAGTGTGGTGTGAAGTACATGCAAGTCTTT 60

QY 1737 TGGCAAGTTCACGCTTATTGTGACAGATGTGAAGCAAGCCTTCATTTTCCCAACAGC 1796
 DB 61 TGGCAAGTTCACGCTTATTGTGACAGATGTGAAGCAAGCCTTCATTTTCCCAACAGC 120

QY 1797 ATTCAAGCGAAAGTCAAGTACAGATGTAGCTATAGCACTAAAGTGGCAATGTGACTGC 1856
 DB 121 ATTCAAGCGAAAGTCAAGTACAGATGTAGCTATAGCACTAAAGTGGCAATGTGACTGC 180
 QY 1857 CAAGATTCAGAAAGTGTGACATTAAGCTATTCCTAGAGGGAGACACAAAGGTTGGCT 1916
 DB 181 CAAGATTCAGAAAGTGTGACATTAAGCTATTCCTAGAGGGAGACACAAAGGTTGGCT 240
 QY 1917 TAAATTTGACACGTGACTGTGAGATCTTTAGTGTGGCTTCATTGACAGAGAGCCGG 1976
 DB 241 TAAATTTGACACGTGACTGTGAGATCTTTAGTGTGGCTTCATTGACAGAGAGCCGG 300
 QY 1977 AAGTCATATCGGGTACAGATGTGGCCACAGAAAGTGGGGGCTCTTCTTAAGCTCTGT 2036
 DB 301 AAGTCATATCGGGTACAGATGTGGCCACAGAAAGTGGGGGCTCTTCTTGAAGCTCTGT 360
 QY 2037 GTCAAGTTCACCTGATATCTTATGTAGATGTGAATGACAAACCTCCAGGCTAGCCAAAGA 2096
 DB 361 GTCAAGTTCACCTGATATCTTATGTAGATGTGAATGACAAACCTCCAGGCTAGCCAAAGA 420
 QY 2097 CTACACGGGCTTGTCTTCTTCCCATATCCCTCAGTGACCTGGAAGTCTCATTTTCCAGGC 2156
 DB 421 CTACACGGGCTTGTCTTCTTCCCATATCCCTCAGTGACCTGGAAGTCTCATTTTCCAGGC 480
 QY 2157 TACTGATGATGATCAGACCTATTTTGGGGTCC 2189
 DB 481 TACTGATGATGATCAGACCTATTTTGGGGATCCC 513

RESULT 10
 AM862312 563 bp mRNA linear EST 19-MAY-2000
 LOCUS AM862312/c
 DEFINITION R4-CT0367-130200-011-b08 CT0367 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM862312
 VERSION AM862312.1 GI:7958007
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 563)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 JOURNAL CONTACT: Simpson A.J.G.
 MEDLINE Laboratory of Cancer Genetics
 COMMENT Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=8&t2=R4-CT0367-130
 200-011-b08&t3=2000-02-13&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 19
 High quality sequence stop: 499.
 Location/Qualifiers
 1..563
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1db="CT0367"
 /dev_stage="Adult"

/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (O.S. Letters Patent Application No. 196 /716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 155 a 113 c 121 g 174 t
ORIGIN

Query Match 14.1%; Score 471.2; DB 10; Length 563;
Best Local Similarity 94.2%; Pred. No. 1.5e-80;

Matches 533; Conservative 0; Mismatches 28; Indels 5; Gaps 4;

QY 960 TATTAGTGAAGTACGCTTGGACCGAAGAAAGATGCAATGTTTTTATGCACT 1019
DB 563 TATTAGTGAAGTACGCTTGGACCGAAGAAAGATGCAATGTTTTTATGCACT 504
QY 1020 TGGAAAGAGTACGCTTGGACCGAAGAAAGATGCAATGTTTTTATGCACT 1079
DB 503 TCGAAGAGTACGCTTGGACCGAAGAAAGATGCAATGTTTTTATGCACT 444
QY 1080 AGATATTATGATATCCACCTACATGCTCCACAGTAACCGTATTGAGTCCAGA 1139
DB 443 AGATATTATGATATCCACCTACATGCTCCACAGTAACCGTATTGAGTCCAGA 384
QY 1140 GAATGAACGACTGGGTACAGTATCGGACCCCTTACTGCAATGACAGGATGAAGAAA 1199
DB 383 GAATGAACGACTGGGTACAGTATCGGACCCCTTACTGCAATGACAGGATGAAGAAA 324
QY 1200 -TATGCAACAGTTTCTTAACTAAGATGTTGAGCAAACTCCCAATCTCCATGG 1258
DB 323 GTACTGCAACAGTTTCTTAACTAAGATGTTGAGCAAACTCCCAATCTCCATGG 264
QY 1259 ATGCACTCTTCTTAACTAAGATGTTGAGCAAACTCCCAATCTCCATGG 1318
DB 263 ATGCACTCTTCTTAACTAAGATGTTGAGCAAACTCCCAATCTCCATGG 204
QY 1319 AGAAGCAAGATATCTCTCACTAAGATGTTGAGCAAACTCCCAATCTCCATGG 1378
DB 203 AGAAGCAAGATATCTCTCACTAAGATGTTGAGCAAACTCCCAATCTCCATGG 144
QY 1379 CCCCTTGTGTTGCAATCAAGTTA-TGATATCATGATGATGATGATGATGATGAT 1437
DB 143 CCCCTTGTGTTGCAATCAAGTTA-TGATATCATGATGATGATGATGATGATGAT 85
QY 1438 AATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1497
DB 84 AATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 27
QY 1498 ACCATCCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1523
DB 26 ACCATCCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1

RESULT 11
AI262603/c 475 bp mRNA linear EST 28-JAN-1999
LOCUS AI262603
DEFINITION gk41h05.x1 NCI CGAP C08 Homo sapiens cDNA clone IMAGE:187157 3'
similar to contains Alu repetitive element, mRNA sequence.
AI262603
VERSION AI262603.1 GI:3870806
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 475)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strauberg, Ph.D.

Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmett-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bio/image/image.html
Insert Length: 838 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 474.
Location/Qualifiers

FEATURES

source

1. 475
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:187157"
/clone_lib="NCI CGAP C08"
/issue_type="adenocarcinoma"
/lab_host="DH10B"

BASE COUNT

166 a 80 c 107 g 122 t
ORIGIN

Query Match 14.1%; Score 470.4; DB 9; Length 475;
Best Local Similarity 99.8%; Pred. No. 2.2e-80;

Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2862 CCCGACCACTGCGCACTAATTTTGTATTTTAAATAGACGGGGTTGGCATTGG 2921
DB 475 CCCGACCACTGCGCACTAATTTTGTATTTTAAATAGACGGGGTTGGCATTGG 416
QY 2922 CCAAGGCTGCTTGAATCTCTGACGTCAGTATCTGCTGCTTGGTCTCCCAATACAG 2981
DB 415 CCAAGGCTGCTTGAATCTCTGACGTCAGTATCTGCTGCTTGGTCTCCCAATACAG 356
QY 2982 GCATGAACCACTGACCACTTATGATATTTGATGATGATGATGATGATGATGAT 3041
DB 355 GCATGAACCACTGACCACTTATGATATTTGATGATGATGATGATGATGATGAT 296
QY 3042 TTTTATTTTTCATGATGATTTTCTCTGCAATGAGCTTACCTGTTGTTTCCC 3101
DB 295 TTTTATTTTTCATGATGATTTTCTCTGCAATGAGCTTACCTGTTGTTTCCC 236
QY 3102 TTTTGGGGCAAGACAGTATTAATATTTCTGATCACTTTTCTTAAAGAGATAT 3161
DB 235 TTTTGGGGCAAGACAGTATTAATATTTCTGATCACTTTTCTTAAAGAGATAT 176
QY 3162 ATCAGTGTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3221
DB 175 ATCAGTGTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 116
QY 3222 TGTCCCTTCATCTTGAATCTCTTGTGATTTTCACTGAATTTCAACATTTGTGAGAA 3281
DB 115 TGTCCCTTCATCTTGAATCTCTTGTGATTTTCACTGAATTTCAACATTTGTGAGAA 56
QY 3282 GAAAAAAGTACGATCAGAAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3333
DB 55 GAAAAAAGTACGATCAGAAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 4

RESULT 12

AA102326 530 bp mRNA linear EST 28-OCT-1996
LOCUS AA102326
DEFINITION 2190E07.r1 Stratiagene colon (#937204) Homo sapiens cDNA clone
IMAGE:511909 5' similar to TR:G854175 G854175 LI-CADHERIN. [1] ;

mRNA sequence.
 AA102326
 AA102326.1 GI:1647016
 EST.
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 530)
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,
 Chisese, S., Dietrich, N., Dubugue, T., Favello, A., Gish, W., Hawkins,
 'M., Hulman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Nardis, E., Moore,
 'B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
 Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevasakis, E.,
 Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
 Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)
 97044478
 COMMENT
 Contact: Wilson RK
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LML; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 347.
 Location/Qualifiers
 1..530
 /organism="Homo sapiens"
 /db_xref="GDB:3844404"
 /db_xref="taxon:9606"
 /clone="IMAGE:511909"
 /clone_lib="Stratagene colon (#937204)"
 /tissue_type="tumor"
 /cell_line="T84 carcinoma cell line"
 /lab_host="SOLR cells (kanamycin resistant)"
 /note="Organ: colon; Vector: pBluescript SK-; Site 1:
 EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
 Oligo dT: T-84 colonic epithelial cell line. Average
 insert size: 1.0 Kb; Uni-ZAP XR Vector; ~5' adaptor
 sequence: 5' GAATTCGCGACGAG 3' ~3' adaptor sequence: 5'
 CTCGAGTTTCTTTTCTTTT 3'."
 BASE COUNT 136 a 116 c 135 g 139 t 4 others
 ORIGIN
 Query Match 14.0%; Score 468.4; DB 9; Length 530;
 Best Local Similarity 97.5%; Pred. No. 5.2e-80;
 Matches 517; Conservative 0; Mismatches 8; Indels 5; Gaps 4;
 QY 1678 TTCAAGCAGAAATCTCTGAGCCCTCTGTTGGTGTGTGAATGCAATGCAAGTCTTTT 1737
 Db 1 TTCAAGCAGAAATCTCTGAGCCCTCTGTTGGTGTGTGAATGCAATGCAAGTCTTTT 60
 QY 1738 GCCAAGTTCAGCTTATTGTGACAGATGTGATGAACACCTCAATTTTCCCAAC-GT 1796
 Db 61 GCCAAGTTCAGCTTATTGTGACAGATGTGATGAACACCTCAATTTTCCCAACNGT 120
 QY 1797 ATTCCAGC-GAAGTCAGTGAGAGATGACTATAGGCACTAAAGTGGGCAATGTGACTG 1855
 Db 121 ATTCCAGCAGAAAGTCAGTGAGAGATGACTATAGGCACTAAAGTGGGCAATGTGACTG 180
 QY 1856 CCAAGATTCAGAAAGTCTGACATATGATCACTGAGAGGAGACACAAAGAGTTGGC 1915
 Db 181 CCAAGATTCAGAAAGTCTGACATATGATCACTGAGAGGAGACAAAGAGTTGGC 240
 QY 1916 TTTAAATTCAGCAC-CTGACTGGTGTGATCTTAGTGTGGCTTCATTGGACAGAGAAGC 1973
 Db 241 TTTAAATTCAGCACCGTGACNTGTGTGATCTTAGTGTGGCTTCATTGGACAGAGAAGC 300
 QY 1974 CGAAGTCATATGCGGTACCAAGTGTGGCACAAGAGTAGGGGGCTCTTCTTAAAGCTC 2033
 .

Db 301 CGAAGTCATATGCGGTACCAAGTGTGGCACAAGAGTAGGGGGCTCTTCTTGAAGCTC 360
 QY 2034 TGTGTGAGATTCACACCTGATCCCTTATGATGATGAATGACACACCTCCAGAGTACCA 2093
 Db 361 TGTGTGAGATTCACACCTGATCCCTTATGATGATGAATGACACACCTCCAGAGTACCA 420
 QY 2094 GAGCTACACGGGCTTGTCTTTCGCAATCCCTCAGTGACACCTGGAAGTCTCA-TTTTGC 2152
 Db 421 GAGCTACACGGGCTTGTCTTTCGCAATCCCTCAGTGACACCTGGAAGTCTCAATTTTGC 480
 QY 2153 AGGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 530
 Db 481 AGGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 530
 RESULT 13
 BM799332
 LOCUS 472 bp mRNA linear EST 05-MAR-2002
 DEFINITION K-EST0083174 S22SNUI6 Homo sapiens cDNA clone S22SNUI6-11-C06 5',
 mRNA sequence.
 ACCESSION BM799332
 VERSION BM799332.1 GI:19147564
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 472)
 Oh, K.J., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
 Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, D.M., Park, H.S., Kim, S. and
 Kim, Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 CONTACT: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribd.re.kr
 Plate: 11 row: C column: 06
 High quality sequence stop: 472.
 Location/Qualifiers
 1..472
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="S22SNUI6-11-C06"
 /clone_lib="S22SNUI6"
 /sex="F"
 /tissue_type="ascites"
 /cell_type="lymphoblast-like"
 /cell_line="SNU-16"
 /lab_host="DH10B"
 /note="Organ: Stomach; Vector: pT73-Pac; Site 1: EcoRI;
 Site 2: NotI. The S22SNUI6 library was constructed as described by
 Soares laboratory and it was constructed as described by
 Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome
 Research 6(9): 791-806. RNA was prepared from harvested
 cells of SNU-16 culture. SNU-16 cell was obtained from
 Korean Cell Line Bank (KCLB). SNU-16 cell was established from
 ascitic fluids of Korean patients by Park J.G. et al.
 (1990), Cancer Res 50: 2773-2780."
 BASE COUNT 102 a 119 c 90 g 161 t
 ORIGIN
 Query Match 14.0%; Score 467.2; DB 14; Length 472;
 Best Local Similarity 99.4%; Pred. No. 9.1e-80;
 Matches 469; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2667 GTGCATTATATTTTAAACAGATATTCCTCTTCTTCTTATATTTGCTAAATATTT 2726
 Db 1 GTGCATTATATTTTAAACAGATATTCCTCTTCTTCTTATATTTGCTAAATATTT 60
 .

Query Match	13.9%;	Score 464.2;	DB 13;	Length 486;
Best Local Similarity	99.4%;	Pred. No. 3.4e-79;		
Matches 466;	Conservative	0;	Mismatches 3;	Indels 0;
			Gaps 0;	
1273 ATCCAAACCTATGCTGGATGTGTACACTTACCTTACCTTGAAGAAGCAAGATCT				1332
486 ATCCAAACCTATGCTGGATGTGTACACTTACCTTACCTTGAAGAAGCAAGATCT				427
1333 CCTCAGTACAACTTAAAGATAGAGGTGCTGACAAAGATTTCAAGACCCCTTTGTTTGTG				1382
426 CCTCAGTACAACTTAAAGATAGAGGTGCTGACAAAGATTTCAAGACCCCTTTGTTTGTG				367
1393 CAATTCACGTTATGTATATCAATGATCAAGATCCCATCTTTGAAAATACAGATTATGGA				1452
366 CAATTCACGTTATGTATATCAATGATCAAGATCCCATCTTTGAAAATACAGATTATGGA				307
1453 AACTGACCTGTGCTGAAGACACAAACATTGGGGTCCACATCTTAAACATTCAGGCCACT				1512
306 AACTGACCTGTGCTGAAGACACAAACATTGGGGTCCACATCTTAAACATTCAGGCCACT				247

OY		1513	GATGCTCATAGACCCATTACTGGAGTGTCTAAATAATCTTGATCAATCATTAAGAAGGAGAC	1572
Dd		246	GATCCTCATATGAGCCATTACTTCGGGAGTTCTAAAATTCGTATCATATACAATAAAGGGAGAC	187
OY		1573	AGTAGGGAAGCGCCCTGGGGGGTGTGCACAGATCCCCATACCACACCGGATATGTATATTT	1632
Dd		186	AGTAGGGAAGCGCCCTGGGGGGTGTGCACAGATCCCATTACCAACCGGATATGTATATTT	127
OY		1633	AAAAAGCCTCTTGATTTTGAAGAAGCAGCGTGTTCACAACATGTGTCTCAAAGCAGAAAT	1692
Dd		126	AAAAAGCCTCTTGATTTTGAAGAAGCAGCGTGTTCACAACATTAATGTTCAAAGCAGAAAT	67
OY		1693	CCTGAGCCTCAGTGTGTGGTGGAAGTATACAAATGCAAGTCTTTGGCA	1741
Dd		66	CCTGAGCCTCAGTGTGTGGTGGAAGTATACAAATGCAAGTCTTTGCCA	18
RESULT 15				
LOCUS	BF334481	477 bp	mRNA	linear EST 22-NOV-2000
DEFINITION	MRI-CT0355-160200-002-g05 CT0355 Homo sapiens cDNA, mRNA sequence.			
ACCESSION	BF334481			
VERSION	BF334481.1	GI:11305333		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 477)			
AUTHORS	Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.G.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A.J., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.			
TITLE	Sequence sequencing of the human transcriptome with ORF expressed			
JOURNAL MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)			
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil			

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MR1ct2-MR1-CT0355-
 160200-002-g05&t3=2000-02-16&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 19
 High quality sequence stop: 41.

FEATURES

SOURCE

1..477
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="CT0355"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site 1: Sma1; Site 2:
 Sma1; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No. 196
 '716 - Ludwig Institute for Cancer Research) profiles
 into the puc 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."

BASE COUNT 160 a 104 c 95 g 118 t
 ORIGIN

Query Match

Best Local Similarity 13.5%; Score 452.8; DB 12; Length 477;
 Matches 460; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1021 GCAAGAGATAGTACGGAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTTAA 1080
 DB 6 GCAATGATAGAGCGAAGAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTTAA 65
 QY 1081 GATATTATATATATATCCACTACATGTCCTGACCAATACCGTATTGAGTCCAGAG 1140
 DB 66 AGATATATATATATATATCCACTACATGTCCTGACCAATACCGTATTGAGTCCAGAG 125
 QY 1141 AATGAAGACTGGGTACAGATGAGGACCCTTACTGCACATGACAGGATGAAGAAAT 1200
 DB 126 AATGAAGACTGGGTACAGATGAGGACCCTTACTGCACATGACAGGATGAAGAAAT 185
 QY 1201 ACTGCCAAGATTTTCTTAACTACAGATTTGTGAGCAAACTCCAAACTTCCATGAT 1260
 DB 186 ACTGCCAAGATTTTCTTAACTACAGATTTGTGAGCAAACTCCAAACTTCCATGAT 245
 QY 1261 GGAATCTTCTTAATCCAACTATGCTGGAATGTTACATTAAGTAAAGTCTTGAAG 1320
 DB 246 GGAATCTTCTTAATCCAACTATGCTGGAATGTTACATTAAGTAAAGTCTTGAAG 305
 QY 1321 AAGCAAGATACCTCTCAGTACAACTTAAGATAGAGGTGCTGACAAAGATTCAAGACC 1380
 DB 306 AAGCAAGATACCTCTCAGTACAACTTAAGATAGAGGTGCTGACAAAGATTCAAGACC 365
 QY 1381 CTTGTTTTTGCAATCAACGTTATGATATCAATGATCAGATCCCATCTTTGAAAAA 1440
 DB 366 CTTGTTTTTGCAATCAACGTTATGATATCAATGATCAGATCCCATCTTTGAAAAA 425
 QY 1441 TCAGATTATGAAACCTGACTTTGCTGAAGACACAAACATTGGGTCCACCA 1492
 DB 426 TCAGATTATGAAACCTGACTTTGCTGAAGACACACAAACATTGGGTCCACCA 477

Search completed: June 22, 2003, 00:52:25
 Job time : 4134 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comphen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2003, 01:08:50 ; Search time 2438 Seconds

(without alignments)
5526.926 Million cell updates/sec

Title: US-10-025-380-1081

Perfect score: 4321

Sequence: 1 MIIQAHLSHLCULMLYATG.....DKGXNVESAQAQSEVPLRS 832

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlj
-Q=/cgn2_1/USPTO/spool/US10025380/runat_17062003_173328_24678/app_query.fasta_1.1031
-DB=EST -QFMT=fastap -SUFPIX=fst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=dot -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10025380 @CGN 1 1 2087 @runat_17062003_173328_24678 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:
1: em_estba:*
2: em_estchum:*
3: em_estcin:*
4: em_estcnu:*
5: em_estcov:*
6: em_estcpl:*
7: em_estcro:*
8: em_hlc:*
9: gb_estc1:*
10: gb_estc2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_ggs:*
18: em_ggs_hum:*
19: em_ggs_inv:*
20: em_ggs_pin:*
21: em_ggs_vit:*
22: em_ggs_fun:*
23: em_ggs_mam:*
24: em_ggs_mus:*
25: em_ggs_other:*
26: em_ggs_pro:*
27: em_ggs_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1227	28.4	1125	14	BM918123	BM918123 AGENCOURT
2	1046	24.2	613	10	AM859603	AM859603 MR1-CT035
3	1032	23.9	625	12	BE696136	BE696136 MR1-CT025
4	955	22.1	573	9	A1347909	A1347909 gp60e04.x
5	884	20.5	873	12	BF582721	BF582721 602094137
6	881	20.4	540	13	BG978653	BG978653 PM1-CN009
7	863	20.0	514	9	A1733765	A1733765 z190f07.y
8	856	19.8	590	10	AM859530	AM859530 MR1-CT035
9	853	19.7	519	10	AM859550	AM859550 MR1-CT035
10	839	19.4	530	9	AA102326	AA102326 z190f07.y
11	819	19.0	657	12	BF584025	BF584025 602096265
12	791	18.3	486	13	BG981361	BG981361 MR1-CT024
13	781	18.1	495	10	AM853021	AM853021 RC1-CN008
14	779	18.0	477	12	BF334481	BF334481 MR1-CT035
15	777	18.0	563	10	AM862312	AM862312 RC4-CT036
16	772	17.9	753	12	BF577684	BF577684 602092567
17	763.5	17.7	620	10	AM377543	AM377543 OV0-CT022
18	750	17.4	455	12	BE696126	BE696126 MR1-CT025
19	734	17.0	507	10	AM753130	AM753130 RC1-CT024
20	726	16.8	551	12	BE808624	BE808624 213876 MA
21	719	16.6	464	10	AM391746	AM391746 CM2-ST022
22	702	16.2	406	10	AM351846	AM351846 RC1-CT019
23	702	16.2	508	9	AA871508	AA871508 vq36c11.x
24	674	15.6	436	13	BG981357	BG981357 CM4-CN008
25	666	15.4	461	10	AM859489	AM859489 MR1-CT035
26	663	15.3	403	9	AA053188	AA053188 z172806.x
27	660.5	15.3	417	10	AM351659	AM351659 QV2-CT014
28	652	15.1	405	13	BG980846	BG980846 CM4-CN008
29	645	14.9	486	9	A1641856	A1641856 vq36c11.y
30	644	14.9	408	12	BF754330	BF754330 IL5-CT051
31	639	14.8	423	9	AA308216	AA308216 EST179051
32	636	14.7	445	13	BG981295	BG981295 CM4-CN008
33	635	14.7	437	10	AM859567	AM859567 MR1-CT035
34	616	14.3	395	10	AM578054	AM578054 RC3-CN001
35	615	14.2	397	10	AM859516	AM859516 MR1-CT035
36	600	13.9	431	13	BG981285	BG981285 CM4-CN008
37	594	13.7	373	13	BG980900	BG980900 CM4-CN008
38	591	13.7	383	13	BG980835	BG980835 CM4-CN008
39	587	13.6	370	12	BF754194	BF754194 IL5-CT051
40	586	13.6	377	10	AM753432	AM753432 MR1-CT025
41	583	13.5	410	13	BG982186	BG982186 CM4-CN008
42	570	13.2	367	13	BG982184	BG982184 CM4-CN008
43	544	12.6	419	10	AM375303	AM375303 QV2-CT014
44	536	12.4	347	10	AM351613	AM351613 QV2-CT014
45	528	12.2	321	13	BG980647	BG980647 CM4-CN008

ALIGNMENTS

RESULT 1
LOCUS BM918123
DEFINITION AGENCOURT 6611470 NIH_MGC_106 Homo sapiens CDNA clone IMAGE:5485497
ACCESSION BM918123
VERSION BM918123.1 GI:19368502
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1125)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgsb@rockefeller.edu
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHC2016 row: h column: 10
High quality sequence stop: 662.
Location/Qualifiers

FEATURES

1.1125
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1ib="NIH MGC 106"
/issue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Blood; Vector: pOT7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC library."

BASE COUNT 259 a 321 c 258 g 286 t 1 others
ORIGIN

Alignment Scores:

Prod. No.: 2.87e-131 Length: 1125
Score: 1227.00 Matches: 240
Percent Similarity: 96.41% Conservative: 2
Best Local Similarity: 95.62% Mismatches: 8
Query Match: 28.40% Indels: 1
DB: 14 Gaps: 0

US-10-025-380-1081 (1-832) x BM918123 (1-1125)

QY 580 AlailegYThrlysvAlGlyAsnvalThralAlysaspProgluglyLeuapilesr 599
DB 1 GCTATGCGCTAAAGGGGCAATGATCTGCCAAGATCCAGAAGTCTGGACATTAAC 60
QY 600 TyrsertleuarglyAspThrarglyTYRpleuylleasphilevalThrglygluile 619
DB 61 TATTCACGTGAGGGAGACACAGAGGTGGCTTAAATTGACACAGCTGGTGAATC 120
QY 620 PheSerValAlaProleuAspArglyAlaGlySerProlyrArgvalGluValAla 639
DB 121 TTATAGTGGCTCATTTGACAGAGAAAGCGGAAGTCAATACGAGTACAGAGTGGG 180
QY 640 ThrGluValGlyGlySerSerleuSerSerValSerGluPheHisleuileuMetasp 659
DB 181 ACAGAGTACGGGGGCTTCTTAAGTCTGTGCAAGTTCACCTGACCTTAATGAT 240
QY 660 ValAsnAspAsnProProArgleuAlaAlaAspTYRThrGlyLeuPhePheCyHisPro 679
DB 241 GTGAATGACAAACCTCCCAAGCTAGCCAGAGACTACACGGGCTTCTTCGCAATCCC 300
QY 680 LeuSerAlaProGlySerleuilePheGluAlaThrAspAspAspGlnHisleuPheArg 699
DB 301 CTAGAGTACCTGAAAGTCTCATTTTGAAGGCTACATATATATATGACACTTAATTTCCG 360
QY 700 GlyProHisPheThrPheSerleuGlySerGlySerleuGlnAsnAspTPGluValaser 719
DB 361 GGTCCCATTTTACATTTTCCCTCGGAGGAGGAAAGCTTACAAAGCACTGGAAAGTTCC 420
QY 720 LysileAsnGlyThrHisAlaArgleuSerThrArgHisThrAspPheGluValargAla 739
DB 421 AAATCAATGATGATCTATGCGGAGCTGTACAGGACACAGACTTGTGAGAGAGGGCG 480

QY 740 TyValValleuileAlargleuAsnAspGlyArgProProleuGluGlyileValser 759
DB 481 TATGTCGTTTATCCGATCAATATGAGGGGTCCGCCACCTTGGAAAGCATTTGTTCT 540
QY 760 LeuProvalThrPheCysSerCyValGlyGlySerCyPheArgProAlaGlyHisgin 779
DB 541 TTACAGTATACATTCCTGCAAGTGTGTGGAGAAAGTGTTCCTCGGACAGTCAACG 600
QY 780 ThrGlyileProThrValGlyMetAlaValGlyleuLeuThrThrleuValile 799
DB 601 ACTGGGATACCACTGTGGGCAAGGAGTGTATCTCTACACCTTCTTGATGAT 660
QY 800 GlyilelleleuAlaValAlaPheleleuArgleuLysAspLysGlyLysAspAsnA 819
DB 661 GGATATATTTTACCGACGTGGCGGCTATCCGCTTAAAGAGACTACGCCAGATATGT 720
QY 819 lgluserAlaGlnAlaSerGluValLysPro 829
DB 721 TGAACGTCTCACGCTCTTGAAGCCACCC 751

RESULT 2
AM859603 613 bp mRNA linear EST 19-MAY-2000
LOCUS
DEFINITION MR1.CT0355-180200-007-e07 CT0355 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM859603
VERSION AM859603.1 GI:7955296
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 613)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Rui Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scrpts/gethtml2.pl?l=kt2=MR1-CT0355-180
200-007-e07&t3=2000-02-18&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 613.
Location/Qualifiers

FEATURES

1.613
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1ib="CT0355"
/dev_stage="Adult"
/note="Organ: Colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 195 a 125 c 134 g 159 t
ORIGIN
Alignment Scores:

Pred. No.: 1.13e-110 Length: 613
 Score: 1046.00 Matches: 202
 Percent Similarity: 100.00% Conservative: 2
 Best Local Similarity: 99.02% Mismatches: 0
 Query Match: 24.21% Indels: 0
 DB: 10 Gaps: 0

US-10-025-380-1081 (1-832) x AW859603 (1-613)

QY 411 LysleuGlnAspThrProGlnIYrAsnLeuThrIleGluValSerAspLysAspPheLys 430
 Db 1 AAGAGAGAGATCTCTCTGAGTACACTTAAACGATAGAGGTGTCTGCAAGATTTCAAG 60
 QY 431 ThrLeuGlyAspPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGlu 450
 Db 61 ACCCTTTGTTTGGCAATCAACGATTATGATCATGATGATGATCCCATCTTTGAA 120
 QY 451 LysSerAspIYrGlyValLeuLeuThrLeuValGluAspThrAsnIleGlySerThrIleLeu 470
 Db 121 AATACAGATTATGAAACCTGACCTTCTGAAAGACCAAACTTGGGTCCACCATCTTA 180
 QY 471 ThrIleGlnAlaThrAspAlaSerGluProPheThrGlySerSerIYrIleLeuYrHis 490
 Db 181 ACCATCCAGCCCACTGATGCTGATGAGCCATTTACTGGAGTTCTAAATTCGTATCAT 240
 QY 491 IleIleGlyValAspSerGluGlyArgLeuGlyValAspThrAspProHisThrAsnThr 510
 Db 241 ATCATAAAGGAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCATACCAACACC 300
 QY 511 GlyIYrValIleIleIleYrProLeuAspPheGluThrAlaValSerAsnIleVal 530
 Db 301 GGATATGTCTAATTAATTAAGCCCTTGATTTGAAACACAGCTGTTTCAACATTTGTG 360
 QY 531 PheIleValGluAsnProGluProLeuValPheGlyValIYrYrAsnAlaSerSerPhe 550
 Db 361 TTCAAGCAAGAAATCTGAGCTCTGATGTTGGTGTGAAGTACAAATGCAAGTCTTTT 420
 QY 551 AlaIYrPheThrLeuIleValIThrAspValAsnGluValProGlnPheSerGlnHisVal 570
 Db 421 GCCAAGTTCAAGCTTATGTCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 571 PheGlnAlaIYrValSerGluAspValAlaIleGlyThrIYrValGlyAsnValThrAla 590
 Db 481 TTCCAAAGCGAAAGTCTGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAG 540
 QY 591 LysAspProGluGlyLeuAspIleSerIYrSerIleAspGlyAspThrArgGlyYrProLeu 610
 Db 541 AAGGATCCAGAAAGTCTGACATPAGCTTATTCATGAGGAGACACAAAGAGGTGGTT 600
 QY 611 LysIleAspHis 614
 Db 601 AAAATTGAACAC 612

RESULT 3
 BE696136/c 625 bp mRNA linear EST 11-SEP-2000
 LOCUS BE696136 MRI-CT0258-040700-002-e10 CT0258 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BE696136
 VERSION BE696136.1 GI:10083296
 ACCESSION
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 625)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.U.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?fl=ec2=MRI-CT0258-040
 700-002-e10&ct3=2000-07-04&ct4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 620.

FEATURES

source
 1..625
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT0258"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
 SmaI; A mini-library was made by cloning products derived
 from ORSTS PCR (U.S. Letters Patent application No. 196
 716 - Ludwig Institute for Cancer Research) profiles
 into the puc 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."

BASE COUNT 143 a 141 c 153 g 188 t
 ORIGIN

Alignment Scores:

Pred. No.: 4.94e-109 Length: 625
 Score: 1032.00 Matches: 202
 Percent Similarity: 98.54% Conservative: 3
 Best Local Similarity: 98.06% Mismatches: 1
 Query Match: 23.88% Indels: 0
 DB: 12 Gaps: 0

US-10-025-380-1081 (1-832) x BE696136 (1-625)

QY 60 PheGluLeuThrGlyGluThrAspAsnIlePheValIleGluArgGluGlyLeuLeuYr 79
 Db 617 TTGAAGCTAAC-GGGAGACAGACACATATTTTGATATGAACGGAGGACTTCTGTAT 559
 QY 80 TyrAsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaLeu 99
 Db 558 TACAAAGAGCCCTTGAGACAGGAAACACAGATCTACACATCTCCAGGTTCAGCCCTG 499
 QY 100 AspAlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValIleAspIle 119
 Db 498 GACGCTAATGATTAATGATGAGGGGTCCAGTCCCTATACCAATGAAGAGGACATC 439
 QY 120 AsnAspAsnArgProThrPheLeuGlnSerIYrGlyGlySerValArgGlnAsnSer 139
 Db 438 AACGAAATCGAACCCAGCTTCTCCAGTCAAAAGTGAAGGCTCAGTAAGGCAACTCT 379
 QY 140 ArgProGlyIYrAspPheLeuYrValAsnAlaThrAspLeuAspAspProAlaThrPro 159
 Db 378 CGCCCAAGAAAGCCCTTCTTGTATGTCATGACACAGACTGGATGATCCGGCACTCC 319
 QY 160 AsnGlyGlnLeuYrIYrGlnIleValIleGlnLeuProMetIleAsnAsnValMetYr 179
 Db 318 AATGGCCAGCTTATTAACAGATGTGATCCAGCTTCCATATCAACAATGTTATGTAC 259
 QY 180 PheGlnIleAsnAsnIYrThrGlyAlaIleSerLeuThrArgGluGlySerGlnGlnLeu 199
 Db 258 TTTCAGATCAACAACAACAGGAGCCATCTCTTACCCGAGAGGATCTCAGAAATG 199
 QY 200 AsnProAlaIYrAsnProSerIYrAsnLeuValIleSerValIYrAspMetGlyGln 219
 Db 198 AATCTGTAAAGATCTTCTATATCTGTGATCTCAGTGAAGACATGGAGGCCAG 139

Qy	220	SergiuAnserPheSeAspThrThrsSerValaPheIleValThnGluAsnIleTrp	239
Db	138	AGTGAATTCTTATGATGATGCCACATCTGTGATATCATAGTACAGAGAAATTTTGG	79
Qy	240	LysAlaProLysProValGluMetValGluAnsSerThrAspProHisPheIleLysIle	259
Db	78	AAAGCACAATACCTGTGGAGTGTGAAAACTCACTGATCTCTCAACCCCATCCAAATC	19
Qy	260	ThrgInValArgTrpAsn	265
Db	18	ACTCAAGTCGGTGGAAAT	1
RESULT 4			
LOCUS	AI347909/c	573 bp	mRNA linear EST 30-DEC-1998
DEFINITION	qp60e04.x1 NCI CGAP C08 Homo sapiens cDNA clone IMAGE:1927422 3'		
ACCESSION	AI347909		
VERSION	AI347909.1	GI:4085115	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 573)		
COMMENT	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
COMMENT	Tumor Gene Index		
COMMENT	Unpublished (1997)		
COMMENT	Contact: Robert Strausberg, Ph.D.		
COMMENT	Email: csgpb-r@mail.nih.gov		
COMMENT	Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.		
COMMENT	Emmert-Buck, M.D., Ph.D.		
COMMENT	cDNA Library Preparation: M. Bento Soares, Ph.D.		
COMMENT	cDNA Library Arrayed by: Greg Lennon, Ph.D.		
COMMENT	DNA Sequencing by: Washington University Genome Sequencing Center		
COMMENT	Clone distribution: NCI-CGAP clone distribution information can be		
COMMENT	found through the I.M.A.G.E. Consortium/ILNI at:		
COMMENT	www.bio.lnl.gov/bbrp/image/image.html		
COMMENT	Seq primer: -40up from Gibco		
COMMENT	High quality sequence scop: 483.		
FEATURES	Location/Qualifiers		
source	1..573		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:1927422"		
	/clone_1ib="NCI CGAP C08"		
	/tissue_type="adenocarcinoma"		
	/lab_host="DH10B"		
	/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a		
	modified polylinker; 1st strand cDNA was prepared from		
	colon adenocarcinoma, and was then primed with a Not I -		
	oligo(dT) primer. Double-stranded cDNA was ligated to Eco		
	RI adaptors (Pharmacia), digested with Not I and cloned		
	into the Not I and Eco RI sites of the modified pT73		
	vector. Library is normalized. Library was constructed by		
	Bento Soares and M. Patricia Bonaldo. "		
BASE COUNT	151 a 155 c 139 g 128 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	3,82e-100	Length:	573
Score:	955.00	Matches:	183
Percent Similarity:	98.94%	Conservative:	3
Best Local Similarity:	97.34%	Mismatches:	1
Query Match:	22.10%	Indels:	1
DB:	9	Gaps:	0
US-10-025-380-1081 (1-832) x AI347909 (1-573)			

[illegible]

Db 418 GACATAGCTATTCTGAGGGGAGACAGAGGTTGGCTAAATTGACACGTGACT 359

Qy 617 GylGluIlePheSerValAlaProLeuAspArgGluAglYserProYlYArgValGln 636

Db 358 GGTGAGATCTTTAGTGTGGCTCCATTGACAGAGAACCCGGAAGTCCATATCGGGTACAA 299

Qy 637 ValValAlaIleThrGluValGlyYserSerLeuSerSerValSerGluPheHisLeuIle 656

Db 298 GTGTGGCCACAGAGAGTGGGGGTCTTCTTAAAGTCTGTGTCAGAGTTCCACCTGATC 239

Qy 657 LeuMetAspValAlaAspAspProProArgLeuAlaIleAspYlYThrGlyLeuPhePhe 676

Db 238 CTTATGATGATGATGATGACCAACCTCCAGGCTACCAAGACATACAGGGCTTTGTTCTTC 179

Qy 677 CysHisProLeuSerAlaProGlySerLeuIlePheGluAlaThrAspAspArgGlnHis 636

Db 178 TGGCATCCCTCATGTCACCTGAGAGTCTCATTTTCAGGCTACTGATGATGATCAGACAC 119

Qy 697 LeuPheArgGlyProHisPheThrPheSerLeuGlyYserGlySerLeuGlnAspAspTyr 716

Db 118 TTATTTGGGGGTCCCTTTTACATTTTCCCTCGGCGAGTGAAGCTTACAAAAGCACTGG 59

Qy 717 GluValSerLysIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAsp 734

Db 58 GAAGTTCCAAATC-AATGGTACTCATGCCGACTGTCCACATGACATGAT 6

RESULT 7

LOCUS A1733765 514 bp mRNA linear EST 14-JUN-1999

DEFINITION z190f07.y5 Stratagene colon (#937204) Homo sapiens cDNA clone

VERSION IMAGE:511909 5' similar to TR:Q12864 Q12864 INTERSTITIAL

KEYWORDS PEPTIDE-ASSOCIATED TRANSPORTER HPT-1. [1] ; mRNA sequence.

ACCESSION A1733765

VERSION A1733765.1 GI:5054878

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 514)

AUTHORS NCI/NIH-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute / National Institute of Dental Research, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Other ESTs: z190f07.x5

Contact: Robert Strauberg, Ph.D.

Email: cgapbs-i@mail.nih.gov

This clone is available royalty-free through LMLT; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

This read is a RESEQUENCE of a previously sequenced human clone

Original clone citation: Washu-Merck EST Project

This read has been verified (found to hit its original self in the correct orientation)

Possible reversed clone: similarity on wrong strand

Seq primer: -40RP from Gdbco

High quality sequence stop: 405.

FEATURES

source

Location/Qualifiers

1. 514

/organism="Homo sapiens"

/db_xref="GDB:384404"

/db_xref="taxon:9606"

/clone="IMAGE:511909"

/clone_lib="Stratagene colon (#937204)"

/rname_type="tumor"

/cell_line="T84 carcinoma cell line"

/lab_host="SOLR cells (kanamycin resistant)"

/note="Organ: colon; Vector: pBluescript SK-; Site 1: EcoRI, Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. T-84 colonic epithelial cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGACGACGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'."

BASE COUNT 135 a 115 c 132 g 133 t

ORIGIN

Alignment Scores:

Pred. No.: 1.57e-89

Score: 863.00

Percent Similarity: 98.254

Best Local Similarity: 98.254

Query Match: 19.974

DB: 9

US-10-025-380-1081 (1-832) x A1733765 (1-514)

Qy 531 PheValAlaGluAspProGluProLeuValPheGlyValYlYThrAsnAlaSerSerPhe 550

Db 2 TTCAAAGCAGAAATCTTGAAGCTTCTAGGTTGGTGTAATACATCAAGTTCTTTT 61

Qy 551 AlaYlePheThrLeuIleValThrAspValAlaGluAlaProGlnPheSerGlnHisVal 570

Db 62 GCCAAGTTCACCTTATTTGACAGATGGAATGAAGCACCTCAATTTTCCCAACACGTA 121

Qy 571 PheGlnAlaIleValSerGluAspValAlaIleGlyThrLysValGlyAsnValThrAla 590

Db 122 TTCCAAGCAGAAAGTCAGTGAAGATGAGTATGAGTCAATGAAGGCGCATGTGACTGCC 181

Qy 591 LysAspProGluGlyLeuAspIleSerYlYSerLeuArgGlyAspThrArgGlyTyrLeu 610

Db 182 AAGATCCAGAAAGTCTGACATTAAGCTTTCTACAGGGGAGACACAGAGGTTGGCTT 241

Qy 611 LysIleAspHisValThrGlyGluIlePheSerValAlaProLeuAspArgGluAglY 630

Db 242 AAAATGACACAGTGAAGTGAAGATCTTATGTGGCTCCATTGACAGAGAGCCGGA 301

Qy 631 SerProTyrArgValGlnValAlaIleThrGluValGlyYserSerLeuSerSerVal 650

Db 302 AGTCATATCGGATTAAGTGTGGCCACAGAGTGGGGGTCTTCTTGAGCTTGtg 361

Qy 651 SerGluPheHisLeuIleLeuMetAspValAlaAspAspProProArgLeuAlaYAsp 670

Db 362 TCAGAGTTCACCTGATCTTATGATGATGATGACACCTCCAGGCTACCAAGAC 421

Qy 671 TyrThrGlyLeuPhePheCysHisProLeuSerAlaProGlySerLeuIlePheGluAla 690

Db 422 TACACGGGCTTGTCTTCTTGCATCCCTCACTGACCCCTGAAGTCTCATTTCCAGGCT 481

Qy 691 ThrAspAspAspGlnHisLeuPheArgGlyPro 701

Db 482 ACTGATGATGATCAGACATATTTTGGATCCA 514

RESULT 8

LOCUS AM859530 590 bp mRNA linear EST 19-MAY-2000

DEFINITION WRI-CT0355-160200-003-e05 CT0355 Homo sapiens cDNA, mRNA sequence.

ACCESSION AM859530

VERSION AM859530.1 GI:7955223

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 590)

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carralho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=et2=MR1-CT0355-160200-004-b02&t3=2000-02-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 22
High quality sequence stop: 577.
Location/Qualifiers

FEATURES

source

1. 590
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_11b="CT0355"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 183 a 127 c 132 g 148 t

ORIGIN

Alignment Scores:

Pred. No.: 1,3e-88 Length: 590
Score: 856.00 Matches: 179
Percent Similarity: 97.30% Conservative: 1
Best Local Similarity: 96.76% Mismatches: 5
Query Match: 19.81% Indels: 3
DB: 10 Gaps: 0

US-10-025-380-1081 (1-832) x AM859550 (1-590)

415 ThrProGlnTyrAsnLeuThrIleGluValSerAspLysAspPheLysThrLeuGlySerPhe 434
19 ACCCCCTGACATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 77
435 ValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLysSerAspTyr 454
78 GTGCAATCAACGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 137
455 GlyAsnLeuThrLeuValGluAspThrAsnIleGlySerThrIleLeuThrIleGlnAla 474
138 GGAAACCTGACTCTTCTGTAAGACACAAACATGGGTCCACCATCTTAACCATCCAGGCC 197
475 ThrAspAlaAspGluProPheThrGlySerSerLysIleLeuThrIleIleValGly 494
198 ACTGATGCTGATGACCATTTACTGGAGTCTTAAATCTGATCATATCATTAAGGGA 257
495 AspSerGluGlyArgLeuGlyValAspThrAspProIleThrAsnThrGlyTyrValIle 514
258 GACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCATACCAACACCGGATATGTCATA 317
515 IleValGlyProLeuAspPheGluThrAlaAlaValSerAsnIleValPheValGlu 534
318 ATTAAAGACCTCTTGAATTTTGAACAGACGCTGTTTCCAACTGTTTCCAAAGCAGAA 377
535 AsnProGluProLeuValPheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThr 554
378 ATCTGAGACCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437
555 LeuIleValThrAspValAsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLys 574
438 CTTATGTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 497
575 ValSerGluAspValAlaIleGlyThrLysValGlyAsnValThrAlaLysAspProGlu 594
498 GTCACTGAGATGTACTATAGCACTAAAGTGGGCCATGTGATGCCAAGGT-CCAGAA 556

QY 595 G1YLeuAspIleSer 599
DB 557 GGCTG-GACATAAGC 570

RESULT 9

AM859550 519 bp mRNA linear EST 19-MAY-2000
LOCUS MR1-CT0355-160200-004-b02 CT0355 Homo sapiens CDNA, mRNA sequence.
DEFINITION AM859550.1 GI:7955243
VERSION EST.
KEYWORDS
SOURCE human.
ORGANISM

REFERENCE
AUTHORS
1 (baee 1 to 519)
Dias Neto,B., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., de Oliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL
MEDLINE
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=et2=MR1-CT0355-160200-004-b02&t3=2000-02-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 81
High quality sequence stop: 517.
Location/Qualifiers

FEATURES

source

1. 519
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_11b="CT0355"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 160 a 111 c 107 g 141 t

ORIGIN

Alignment Scores:
Pred. No.: 2,33e-88 Length: 519
Score: 853.00 Matches: 166
Percent Similarity: 98.26% Conservative: 3
Best Local Similarity: 96.51% Mismatches: 3
Query Match: 19.74% Indels: 0
DB: 10 Gaps: 0

US-10-025-380-1081 (1-832) x AM859550 (1-519)

415 ThrProGlnTyrAsnLeuThrIleGluValSerAspLysAspPheLysThrLeuGlySerPhe 434
DB 3 ACTCTGATATTAACCAACATTAAGAGTGTGACACAGATTAACAAACCTTTGTTT 62
QY 435 ValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLysSerAspTyr 454

[illegible]

BASE COUNT		136 a	116 c	135 g	139 t	4 others
ORIGIN						
/lab host="SOLR cells (kanamycin resistant)"						
/note="Organ: colon; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. T-84 colonic epithelial cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGACGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3'"						
US-10-025-380-1081 (1-832) x AA102326 (1-530)						
Alignment Scores:		1,02e-86	Length:	530		
Prd. No.:	839.00	Matches:	174			
Score:	96.67%	Conservative:	0			
Percent Similarity:	96.67%	Mismatches:	1			
Best Local Similarity:	19.42%	Indels:	5			
Query Match:	9	Gaps:	0			
DB:						
Oy	531	PhelysAlaGluAsnProGluProLeuValPheGlyValIysTyrAsnAlaSerSerPhe	550			
Db	1	TTCAAAACAGAAATCCTGAGCCCTCTGCTTTGGTGTGAAGACAAATGCAAGTCTCTTT	60			
Oy	551	AlAlaySerPheIleuIleValIThrAspValAsnGluAlaProGlnPheSerGlnHis-Va	570			
Db	61	GCCAAAGTTCACGCTTATTTGACAGAGTGAATGAAGACACCTCAATTTTCCCAACACNGT	120			
Oy	570	IPheGlnAla-LysValSerGluAspValAlaIleGlyThrIlyValGlyAsnValThra	590			
Db	121	ATTCACAGCGAAGAAAGTCAGTGAGGATGTACTGATACCTAAAGTGGGCAATGTGACTG	180			
Oy	590	IAlYsAspProGluGlyLeuAspIleSerTyrSerLeuArgGlyAspThrArgGlyTrrp	610			
Db	181	CCAAGATTCAGAAAGGTCTGGACATTAAGCTAATTCACAGAGGGGAGACACAAAGAGTTGGC	240			
Oy	610	eulysIleAspHis-ValIThr-GlyGluIlePheSerValAlaProLeuAspArgGluAl	629			
Db	241	TTAAATTTGACCAACCGGACGTCGATGGTGTGAGATCTTTAGTGTGCTTCATTGGACAGAGAC	300			
Oy	629	aGlySerProTyrArgValGlnValValAlaIleGluValGlyIySerSerLeuSerSe	649			
Db	301	CGAAGATCCATATCGGGTTCACAGTGGTGGCCACAGAAAGTGGGGGGCTTCCTTGAAGCTC	360			
Oy	649	rValSerGluPheHisLeuIleLeuMetAspValAsnAspAsnProProArgLeuAlaIy	669			
Db	361	TGTGTACAGAGTTCACCTGATCCTTATGATGTGAATGAACACCTCCACAGGCTAGCCAA	420			
Oy	669	aAspTyr-ThrGlyLeuPhePheCysHisProlSerSerAlaProGlySerLeuIle-PheG	689			
Db	421	GGACTACACCGGGCTGTGCTTCGCAATCCCTCAGTGACACCTGGAGGTCTCATTTTTCG	480			
Oy	689	IuAlaThrAspAspAspGlnHisIleuPheArgIyProHisPheThrPhe	705			
Db	481	AGGCTACTGATGATGATCAGCAGCTATTATTGGGGGAGGCCAATTTTACATTT	530			
RESULT 11						
LOCUS	BF584025	657 bp	mRNA	linear	EST 12-DEC-2000	
DEFINITION	6020962265P1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4216216 5',					
ACCESSION	BF584025					
VERSION	BF584025.1	GI:11657743				
KEYWORDS	EST.					
SOURCE	house mouse.					
ORGANISM	Mus musculus					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.					
TITLE	NIH-MGC http://mgc.nci.nih.gov/					
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)					
COMMENT	Unpublished (1999)					
	Contact: Robert Strausberg, Ph.D.					

Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1AM9793 row: a column: 17
 High quality sequence stop: 655.

FEATURES

source

1. 657
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4216216"
 /clone_1ib="NCI CGAP CO24"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: colon; Vector: PCMV-SPORE6; Site_1: Nci1;
 Site_2: Sal1; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.6 kb. Constructed by Life
 Technologies. Note: this is a NCI CGAP Library."
 BASE COUNT 166 a 158 c 176 g 157 t
 ORIGIN

Alignment Scores:

Pred. No.: 3,12e-84 Length: 657
 Score: 819.00 Matches: 164
 Percent Similarity: 84.02% Conservative: 20
 Best Local Similarity: 74.89% Mismatches: 35
 Query Match: 18.95% Indels: 2
 DB: Gaps: 0

US-10-025-380-1081 (1-832) x BF564025 (1-657)

539 LeuValPheGlyValIleYrValAsnAlaSerSerPheAlaIlePheThrLeuIleValThr 558
 2 CTGGTAAAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
 559 AspValAsnGluAlaProGlnPheSerGlnIleValPheGlnAlaValSerGluAsp 578
 61 GATGTAAATGAAGTGGCTGT 120
 579 ValAlaIleGlyThrIleValGlyAsnValThrAlaValAspProGluGlyLeuAspIle 598
 121 GCGCTGTGGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 599 SerTyrSerLeuArgGlyAspThrArgGlyTyrIleValIleAspHisValThrGlyGlu 618
 181 AGTTATTCAGTAAAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 619 IlePheSerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAla 638
 241 AATATTAGTGTCTGT 300
 639 AlaThrGluValGlyIleSerSerLeuSerSerValSerGluPheHisLeuIleLeuMet 658
 301 GCCACGTGAAGTGGT 360
 659 AspValAsnAspAsnProProArgLeuAlaIleAspTyrThrGlyLeuPhePheCysHis 678
 361 GATGTAAATGACAAACCCCTGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 420
 679 ProLeuSerAlaProGlySerLeuIlePheGluAlaThrAspAspArgGlnIleLeuPhe 698
 421 CCCCTAGTGGCCCTGGAGCTCATCTTTGAGTCACTGACATGACATGACATGACATGAC 480
 699 ArgGlyProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluVal 718
 481 CCGAGGCCCAAGTTTATTCGCGCTTGGCAGAGAGGCTTCAAGAGTGACTGGAGAGTT 540
 719 SerIleValAsnGlyThrHisAlaIleArgLeuSerThrArgHisThrAspPheGluGluArg 738

DB 541 TCAAAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 739 AlaTyrValIleValLeuIleArgIleAsnAspGlyIleArgProProLeuGluGlyIle 757
 DB 601 GTTTATTAACATTTCCATTCGATCATCATCATCATCATCATCATCATCATCATCATCAT 656

RESULT 12
 BG981361/c 486 bp mRNA linear EST 12-JUN-2001
 LOCUS CM4-CN0089-070201-746-g10 CN0089 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BG981361
 ACCESSION BG981361 GI:14384096
 VERSION BG981361.1
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 486)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Britones, M. R.,
 Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, P. F.,
 Goldman, G. H., Carvalho, A. F., Matsukuma, A., Baia, G. S., Simpson, D. H.,
 Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare,
 M. O., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and
 Simpson, A. J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 CONTACT: Simpson A.J.G.
 LABORATORY: Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=CM4&cl2=CM4-CN0089-
 070201-746-g10&cl3=2001-02-07&cl4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 3
 High quality sequence stop: 482.
 Location/Qualifiers

FEATURES

source

1. 486
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1ib="CN0089"
 /dev_stage="Adult"
 /note="Organ: colon normal; Vector: puc18; Site_1: Sma1;
 Site_2: Sma1; A mini-library was made by cloning products
 derived from ORSTES PCR (O.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 134 a 95 c 104 g 153 t
 ORIGIN

Alignment Scores:

Pred. No.: 3.36e-81 Length: 486
 Score: 791.00 Matches: 154
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.35% Mismatches: 0
 Query Match: 18.31% Indels: 0
 DB: Gaps: 0

US-10-025-380-1081 (1-832) x BG981361 (1-486)

336 IleGlnThrTyrAlaGlyMetLeuGlnIleValAlaValGlnIleValGlnIleValThr 415
 486 ATCCAAACCTATGCTGGAATGTTACAGTTAGCTAAACAGTCTTGAAGACAGATGACT 427

QY 416 ProGlnTyrAsnLeuThrIleGluValSerAspLysAspPheLysThrLeuGlyCysPheVal 435
 DB 426 CCTCAGACCACTTAAGATGAGGTCTGACCAAGATTTCAGACCTTTGTTTGTG 367
 QY 436 GlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLysSerAspTyrGly 455
 DB 366 CAATCAACGTTATGATATCAATGATCAGATCCCATCTTTGAAAATCAGATTATGCA 307
 QY 456 AsnLeuThrIleuValaGluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThr 475
 DB 306 AACCTGACTCTTCTGTAAGACACAAACATGGGTGCACCATTTAAACATCCAGGCCACT 247
 QY 476 AspAlaAspGluProPheThrGlySerSerLysIleLeuTyrHisIleIleLysGlyAsp 495
 DB 246 GATGCTGATAGACCATTTACTGGAGTTCTTAAATTTCTGATCATATCAAAAGGAGAC 187
 QY 496 SerGlnGluValArgLeuGlyValaAspThrAspProHisThrAsnThrGlyTyrValIleIle 515
 DB 186 AGTAGAGGACGCTGGGGGTGACACAGATCCCATCCAAACCGGATATGTCATATTT 127
 QY 516 LysLeuProLeuAspPheGluThrAlaAlaValSerAsnIleValaPheLysAlaGluAsn 535
 DB 126 AAAAAGCCTCTTGAATTTTGAACAGCAGCTGTTCCAACTTATGTTCAAGAGAAAT 67
 QY 536 ProGluProLeuValaPheGlyValaLysTyrAsnAlaSerSerPhe 550
 DB 66 CCTGAGCCTTACTGTTTGTGTTGATGATCAATGCAAGTCTTTT 22
 RESULT 13
 AM853021/c 495 bp mRNA linear EST 19-MAY-2000
 LOCUS R1-CT0249-170200-025-b11 CT0249 Homo sapiens cDNA, mRNA sequence.
 DEFINITION AM853021
 VERSION AM853021.1 GI:7948538
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 495)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 TITLE Contact: Simpson A.J.G.
 JOURNAL Laboratory of Cancer Genetics
 MEDLINE Ludwig Institute for Cancer Research
 COMMENT Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPER/P/LICK Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=RC1-CT0249-170
 200-025-b11&f3=2000-02-17&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 13
 High quality sequence stop: 495.
 Location/Qualifiers
 1..495
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="CT0249"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
 SmaI; A mini-library was made by cloning products derived

from ORESTES PCR (U.S. Letters Patent application No. 196
 716 - Ludwig Institute for Cancer Research) profiles
 into the pUC 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."

BASE COUNT 121 a 108 c 124 g 142 t
 ORIGIN

Alignment Scores:
 Pred. No.: 5,04e-80 Length: 495
 Score: 781.00 Matches: 133
 Percent Similarity: 96.23% Conservative: 0
 Best Local Similarity: 96.23% Mismatches: 6
 Query Match: 18.07% Indels: 1
 DB: 10 Gaps: 0

US-10-025-380-1081 (1-832) x AM853021 (1-495)

QY 6 HisLeuHisSerLeuGlyLeuLeuMetLeuTyrLeuAlaThrGlyTyrGlyGlnGluGly 25
 DB 493 CATCTTCACTCCCTGCTGCTCTTTATGCTTATTTGGCACTGATATGCCAAGAGGGG 434
 QY 26 LysPheSerGlyProLeuLysProMetThrPheSerIleTyrGlyGlnGluProSer 45
 DB 433 AATTTTAGTGAGACCCCTGAACCCATGACATTTCTATTATTAAGAGCCCAAGACCACT 374
 QY 46 GlnIleIlePheGlnPheLysAlaAsnProAlaValaThrPheGluLeuThrGlyGlu 65
 DB 373 CAATTAATTCACAGTTAAGCCCAATCTCTGCTGATCTTGAATTAAGTGGGAG 314
 QY 66 ThrAspAsnIlePheValaIleGluArgGluGlyLeuLeuTyrTyrAsnArgAlaLeuAsp 85
 DB 313 ACAGACAACTATTTGTGTATGATAGACGGAGGACTTCTGTATTAACAGACCTTGGAC 254
 QY 86 ArgGluThrArgSerThrHisAsnLeuGlnValaAlaLeuAspAlaAsnGlyIleIle 105
 DB 253 AGGAAACAAAGATCTACTACATCTCCAGTTGGACGCTCGAGCCTTAATGAAATTATA 194
 QY 106 ValGluGlyProValProIleThrIleGluValaLysAspIleAsnAspAsnArgProThr 125
 DB 193 GTGAGAGGTCCAGTCCCTTATCCATAGAAAGGAGACATCAACACATGACACCAAG 134
 QY 126 PheLeuGlnSerLysTyrGlyGlySerValaArgGlnAsnSerArgProGlyLysProPhe 145
 DB 133 TTTCCTCACTCAAGATGACAGAGGCTCAGTAAGGAGAACTCTCGCCAGAAAGCCCTTC 74
 QY 146 LeuTyrValaAsnAlaThrAspLeuAspProAlaThrProAsnGlyGlnLeuTyr 164
 DB 73 TTGTATGTCAAATGCCACAGACCTGATGATCCGCC-ACCTCCCAATGCCAGCTTATAC 18
 RESULT 14
 BF334481 477 bp mRNA linear EST 22-NOV-2000
 LOCUS MR1-CT0355-160200-002-g05 CT0355 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF334481
 ACCESSION BF334481.1 GI:11305333
 VERSION
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 477)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 TITLE Contact: Simpson A.J.G.
 JOURNAL Laboratory of Cancer Genetics
 MEDLINE Ludwig Institute for Cancer Research
 COMMENT Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPER/P/LICK Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=RC1-CT0249-170
 200-025-b11&f3=2000-02-17&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 13
 High quality sequence stop: 495.
 Location/Qualifiers
 1..495
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="CT0249"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
 SmaI; A mini-library was made by cloning products derived

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MRI&ct=MRI-CT0355-
160200-002-905&ct3=2000-02-16&ct4=1)
Seq primer: puc 18 forward
High quality sequence start: 19
High quality sequence stop: 41.
Location/Qualifiers

FEATURES

source

```
1. 477
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="CT0355"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
```

BASE COUNT 160 a 104 c 95 g 118 t

ORIGIN

Alignment Scores:

```
Pred. No.: 8.09e-80 Length: 477
Score: 779.00 Matches: 151
Percent Similarity: 96.82% Conservative: 1
Best Local Similarity: 96.18% Mismatches: 5
Query Match: 18.03% Indels: 0
DB: 12 Gaps: 0
```

US-10-025-380-1081 (1-832) x BF334481 (1-477)

```
QY 312 AAlaAAspGluTyrGlyPProLeuSerTyrProLeuGluLeuHsVallyVally 331
DB 6 GCATGAGATGAGAGACGAAACCACTTCAATCCCTGGAAATACATGTAAGAGATAA 65
QY 332 AspIleAsnAspAsnProProThrCysProSerProValThrValPheGluValGlnGlu 351
DB 66 AGATATATATGATATATCACTACATGTCCTCCACCGATACCGTATTGGAGTCCAGAG 125
QY 352 AsnGluArgLeuGluGlnSerIleGlyThrLeuThrAlaHsAspArgAspGluGluAsn 371
DB 126 AATGAAAGCACTGGGTAAAGATATCGGACCTTACATGCAATGACATGACGGATGAAGAAAT 185
QY 372 ThrAlaAsnSerPheLeuAsnTyrArgIleValGlnGlnThrProLysLeuProMetAsp 391
DB 186 ACTGCCAAGATTTTCTTAACTACAGATTTGTGACAAACTCCAACTTCCCATGAT 245
QY 392 GlyLeuPheLeuIleGlnThrTyrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLys 411
DB 246 GGAATCTTCTCTAATCCAAACCTAAGCTGAGATGTACAGTAAACAGTCTTGAAG 305
QY 412 LysGlnAspThrProGlnTyrAsnLeuThrIleGluValSerAspLysAspPheLysThr 431
DB 306 AAGCAAGATATCTCTCAGTAAACCTTAAGATAGAGTGTCTGACAAAGATTTCAGAGC 365
QY 432 LeuGlyPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLys 451
DB 366 CTTGTTTGTGTGCAATCAACGTTATGATATCAATGATCAATGCCATCTTTGAAAAA 425
QY 452 SerAspTyrGlyAsnLeuThrLeuAlaGlyAspThrAsnIleGlySerThr 468
DB 426 TCAGATTATGAAACCTGACTTGTCTGAAGACACACATGGGTCCACC 476
```

RESULT 15

AM862312/c
LOCUS AM862312 563 bp mRNA linear EST 19-MAY-2000
DEFINITION RC4-CT0367-130200-011-b08 CT0367 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM862312
VERSION AM862312.1 GI:7958007
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 563)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Birstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=RC4-CT0367-130
200-011-b08&ct3=2000-02-13&ct4=1)
Seq primer: puc 18 forward
High quality sequence start: 19
High quality sequence stop: 499.
Location/Qualifiers

FEATURES

source

```
1. 563
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="CT0367"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
```

BASE COUNT 155 a 113 c 121 g 174 t

ORIGIN

```
Alignment Scores:
Pred. No.: 1.83e-79 Length: 563
Score: 777.00 Matches: 163
Percent Similarity: 92.97% Conservative: 9
Best Local Similarity: 88.11% Mismatches: 11
Query Match: 17.98% Indels: 4
DB: 10 Gaps: 0
```

US-10-025-380-1081 (1-832) x AM862312 (1-563)

```
QY 292 IleTyrValThrGlnPProLeuAspArgGluGlyAspAlaTyrValPheTyrAlaVal 311
DB 562 ATTTCAGTACCCACCCCTTGGACGAGAGATGATATATTTTATCAATT 503
QY 312 AAlaAAspGluTyrGlyPProLeuSerTyrPProLeuGluIleHsVallyVally 331
DB 502 CCAAGAGATGATGAGTAAACCACTTCAATTCCTCGGAATTCATGTAAGATTA 443
QY 332 AspIleAsnAspAsnProProThrCysProSerProValThrValPheGluValGlnGlu 351
DB 442 GAGATTATGATTAATCCACCTAATGATGTCGTCACAGTAACCGATTTTGAGGTCCAGAG 383
```

```
QY 352 AaenGIuArgIeuGIyAsnSerIleGIyThrLeuThrAlaHisAspArgAspGIu-As 371
Db 382 AATGTACGACTGGGTTCAGTATCGGAGCCCTTACTGCACATGACAGGATGAGAAAG 323
QY 371 nThrAlaAsnSerPheLeuAsnTyrArgIleValGIuGIuThrProIyLeuPheMetAs 391
Db 322 TACTGCCAACAGTTTCTATACACAGATTGTGAGCAAACTCCAAACTTCCATGGA 263
QY 391 pGIyLeuPheLeuIleGIuThrTyrAlaGIyMetLeuGIuLeuAlaLysGIuSerLeuIy 411
Db 262 TGGACTCTTCTATACCGAACCTATGCTGGAATGTACAGTATACAGTCTTGAA 203
QY 411 sLyGIuAspThrProGIuThrAsnLeuThrIleGIuValSerAspLysAspPheIyStH 431
Db 202 GAAGCAAGATACCTCTCAGTACAACTTAACGATAGAGGTGTGACAAAGATTTCAGAC 143
QY 431 rLeuCySPheValGIuIleAsnVal-IleAspIleAsnAspGIuIleProIlePheGIuIu 451
Db 142 CCTTTGTTTGTGCAAAATCAACGTTAGTATATCAATGATCAGATCCCATCTTGTAT 83
QY 451 ySserAspTyrGIyAsnLeuThrLeuAlaGIuAspThrAsnIleGIySerThrIleLeuT 471
Db 82 A-TCAGATATAGTAACCTGACTCTTGTGAGACACACACATTGG--TCCACCATCGTAA 26
QY 471 hrlIeGIuAla 474
Db 25 CCATCTTGCA 15
```

Search completed: June 22, 2003, 03:20:27
Job time : 2447 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2003, 00:52:29 ; Search time 63 seconds
(without alignments)
2721.130 Million cell updates/sec

Title: US-10-025-380-1081

Perfect score: 4321

Sequence: 1 MIQAHLSLCILMYLATG.....DKGKNVESQAEVPLNS 832

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2332	54.0	578	11	Q63423
2	1427.5	33.0	868	13	Q90X63
3	709.5	16.4	893	13	Q80VQ7
4	699	16.2	922	13	P79883
5	677	15.7	783	13	Q90275
6	671.5	15.5	839	4	Q9HA29
7	652	15.1	713	4	Q8T8X3
8	652	15.1	714	11	Q8VDC4
9	651.5	15.1	814	6	Q77704
10	650	15.0	714	11	Q8R490
11	628	14.5	4589	11	Q9WU10
12	624	14.4	4587	11	Q90XA3
13	611	14.1	4555	11	Q8R508
14	601.5	13.9	3579	5	Q9V5N8
15	591.5	13.7	4643	5	Q9VW71
16	576	13.3	3014	4	Q9NY06

17	572.5	13.2	901	4	Q9U117	Q9U117 homo sapien
18	571	13.2	3034	11	Q35161	Q35161 mus musculus
19	570.5	13.2	840	4	Q9HB00	Q9HB00 homo sapien
20	570.5	13.2	894	4	Q9HB01	Q9HB01 homo sapien
21	564	13.0	3301	11	Q91Z10	Q91Z10 mus musculus
22	561.5	13.0	796	4	Q96C29	Q96C29 homo sapien
23	558	12.9	3313	11	Q88278	Q88278 rattus norv
24	557.5	12.9	3312	4	Q9NY07	Q9NY07 homo sapien
25	556.5	12.9	788	11	Q8V168	Q8V168 mus musculus
26	552	12.8	3503	5	Q24292	Q24292 drosophila
27	545.5	12.6	801	11	Q9Z0M3	Q9Z0M3 mus musculus
28	543.5	12.6	693	4	Q15066	Q15066 homo sapien
29	543.5	12.6	792	13	Q9DF80	Q9DF80 xenopus lae
30	543.5	12.6	798	13	Q8GCH3	Q8GCH3 gallus gall
31	541.5	12.5	792	13	Q9DF81	Q9DF81 xenopus lae
32	538	12.5	864	13	Q90Z37	Q90Z37 brachydanto
33	530	12.3	2923	4	Q9HC04	Q9HC04 homo sapien
34	528	12.3	2809	5	Q61230	Q61230 lytechinus
35	521	12.1	4349	4	Q9NY08	Q9NY08 homo sapien
36	518.5	12.0	4351	11	Q88277	Q88277 rattus norv
37	518	12.0	782	6	Q8N8W5	Q8N8W5 sus scrofa
38	515.5	11.9	910	5	Q95YK9	Q95YK9 citra savig
39	514.5	11.9	503	4	Q9NTY1	Q9NTY1 homo sapien
40	508	11.8	2920	11	Q9ROM0	Q9ROM0 mus musculus
41	496.5	11.5	790	13	Q91838	Q91838 xenopus lae
42	494	11.4	794	13	Q93264	Q93264 xenopus lae
43	489	11.3	828	4	Q9U099	Q9U099 homo sapien
44	489	11.3	1783	5	Q9VDB6	Q9VDB6 drosophila
45	488.5	11.3	821	4	Q9U118	Q9U118 homo sapien

ALIGNMENTS

RESULT 1

Q63423 PRELIMINARY; PRT; 578 AA.

AC Q63423;
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBLrel. 21, Last annotation update)
DE Proton-driven peptide transporter (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=NISTAR; TISSUE=SMALL INTESTINE;
RX MEDLINE=96067558; PubMed=7488096;
RA Erickson R.H., Gunn J.R. Jr., Lindstrom M.M., McKean D., Kim Y.S.;
RT "Regional expression and dietary regulation of rat small intestinal
peptide and amino acid transporter mRNAs."
RL Biochem. Biophys. Res. Commun. 216:249-257(1995).
CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
DR EMBL: I46874; AAC2077.1;
DR HSSP: P15116; INCI.
DR InterPro: IPR002126; Cadherin.
DR Pfam: PF00028; cadherin; 6.
DR SMART: SM00112; CA; 6.
DR PROSITE: PS00232; CADHERIN_1; 1.
DR PROSITE: PS50268; CADHERIN_2; 6.
KW Calcium-binding; Cell adhesion; Glycoprotein.
FT NON_TER
FT NON_TER
SQ SEQUENCE 578 AA; 64517 MW; 5644574BC6D14F80 CRC64;

Query Match 54.0%; Score 2332; DB 11; Length 578;
Best Local Similarity 77.3%; Pred No. 4,3e-145;

Matches 447; Conservative 55; Mismatches 76; Indels 0; Gaps 0;

QY 62 LTGETDNIFVIERGLLYNRALDRETRSTNQLQVALDANGIIVEGPVITREVDIND 121
||||| |::|||:| |||||:|::|||:| |||||

```

Db      1  LTGETDGLIKIEKQGLIYHARVLDRETRAVHHLRLALMSQGAIVDGPVPIIEVDIND 60
Qy      122  NRPTFLOSKTEGVRONSRPGKPEFLYNATDLDPPATNGOLYYQIYQIQLPMINNNVYFQ 181
Db      61  NRPTFLOSKTEGVRONSRPGKPEFLYNATDLDPPATNGOLYYQIYQIQLPMINNNVYFQ 120
Qy      182  INNKTAISLTRESGOELNPAKNPSYVLVI SVKMGQSGNSFSDTSSVLIITENIMKA 241
Db      121  IDNKTAISLTRESGOELNPAKNPSYVLVI SVKMGQSGNSFSDTSSVLIITENIMKA 180
Qy      242  PKREVENENSTDPHPKITQVRWMDPGAOYSLVDEKELPRPPSIDEGDIYVTOPLDRE 301
Db      181  PEVEIERENLTDPHPKITQVRWMDPGAOYSLVDEKELPRPPSIDEGDIYVTOPLDRE 240
Qy      302  EKQAYFYAVAKDEYKPELSTYPLEIHYKVDINDNPTCSPTVFEVQENRGLNSIGT 361
Db      241  EKQSHVFAACQKENGKPLAYPLEIRVKVIDINDNPTCSPTVFEVQENRGLNSIGT 300
Qy      362  LTAHDREENTANSFLAYRIVEOTPKLPMQGLFIQYAGMLQAKOSLKKODTPQVNL 421
Db      301  PAHMDMEANNINSLMYRLVDOTPKVPSDELFLIDYGGKVLGKSLKKODSPQYNLT 360
Qy      422  IEVSDKDFKLCFQVQINVIDINDQIPIFEKSDYGNLTAEPTNIGSTILTIOATDADPE 481
Db      361  VEVSIDDFKLCFQVQINVIDINDQIPIFEKSDYGNLTAEPTNIGSTILTIOATDADPE 420
Qy      482  TGSSKILYHIKQDSBGLGVTDPHNTGYIILKKLPDETAANSIVKAEPEPLV 541
Db      421  TGSSKILYHIKQDSBGLGVTDPHNTGYIILKKLPDETAANSIVKAEPEPLV 480
Qy      542  GVKNASSFAKFTLIVDVNEAPQSOHVFOAKYSEVAIGTKVGNATADPEGLDSSYS 601
Db      481  GIEKNASSFAKFTLIVDVNEAPQSOHVFOAKYSEVAIGTKVGNATADPEGLDSSYS 540
Qy      602  LRQDTRGMLKIDHYTGEIFSVAPLDRAGSEYRQVVA 639
Db      541  LKONKRGMLKIDHYTGEIFSVAPLDRAGSEYRQVVA 578

RESULT 2
ID 090X63  PRELIMINARY;  PRT;  868 AA.
AC 090X63;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Cadherin-17.
GN CDH17.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_Taxid=7955;
RN [1]
RA HOFSTEDT J., RAMACHANDRAN A., REUTER K., LAVALLIE E.,
RA COLLINS-RACE L., CROSTER K., CROSTER P. ST.,
RT "zebrafish cadherin-17 is regulated for pro-neuronic duct development.",
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
DR EMBL; AF280988; AAL29444.1;--
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 6.
DR PROSITE; PS00232; CADHERIN_1; UNKNOWN_2.
DR PROSITE; PS50268; CADHERIN_2; 6.
KM Calcium-binding; Cell adhesion; Glycoprotein.
SQ SEQUENCE 868 AA; 96677 MW; 40C30844C836E990 CRC64;

Query Match 33.0%; Score 1427.5; DB 13; Length 868;
Best Local Similarity 36.5%; Pred. No. 2.5e-85;
Matches 320; Conservative 150; Mismatches 330; Indels 77; Gaps 16;
Qy 13 LMLYLIANGYG--QGKSGSLPKMTFTSYEGQSFQIIFQF--KANPRAVTFELNGETDNI 69

```

```

Db      9  LTLVLSIGHGIDLEBK--KGFLIDTVLDVEAPVPAFAFKFTSAVEDVSSYRSGETEDK 67
Qy      70  FYIEREGLLYNNRAIDRETRSTHNLQVALMDANGIIVEEPVITIEVKDINDNRPFTLOS 129
Db      68  IRISDGMVLYLQPLEMSDEKGHLLIEALSBDGKLDGPAAVYQVQIDVNNPVSFS 127
Qy      130  KYEGSVRONSRPGKPEFLYNATDLDPPATNGOLYYQIYQIQLPMINNNVYFQIINKTGA 189
Db      128  QIGSIRHSIPAGVPVQVPAADDDPNTENTQLRISYNOQPIVQGTFFPGINPNNGQ 187
Qy      190  SLTREGSOEL----- 199
Db      188  FTTEBGAFLKARPSTYTSRGEVRGSDVLKKKFEEDYCIPKNNIALENNPFYKVERAER 247
Qy      200  ---NPAKNPSYVLVI SVKMGQSGNSFSDTSSVLIITENIMKA PKPMBMENSNDPP 256
Db      248  RTVNVLIQDDPDVALIYAEDLGNAVNSLSTTRVNTAILQNLVWSFGPTIRENTEDEAP 307
Qy      257  IKITQVRWMDPGAOYSLVDEKELPRPPSIDEGDIYVTOPLDREKQAYFYAVAKDEY 316
Db      308  MYLATRANNPTRALYRLQKEL--SPTTINDGDIYVTOPLDREKEMVILVIAEDQ 366
Qy      317  GKPLSYPLEIHYKVDINDNPTCSPTVFEVQENRGLNSIGTLTAHDREENTANSF 376
Db      367  GVELEKPMELPVLYQDENPNPWCDE--ALFEVQEKEPVNSIGHLPADHNDREGTLSA 424
Qy      377  LMRIVEOTPKLPMQGLFIQYAGMLQAKOSLKKODTPQVNLTEVSDKDFKLCFQ 436
Db      425  LVTTLRSQTPKRSDFMSIDPNTGKIVANONFOKKQVQVELFEVVDQVFFTCQAI 484
Qy      437  INVINDINDQIPIFEKSDYGNLTAEPTNIGSTILTIOATDADPEFGSSKILYHIKQDS 496
Db      485  IKVIDINDQIPIFEKSDYGNLTAEPTNIGSTILTIOATDADPEFGSSKILYHIKQDS 544
Qy      497  EGRIGVTDPHNTGYIILKKLPDETAANSIVKAEPEPLVFGVKNASSFAKFTLI 556
Db      545  QNLFALVEDEBERGEGRVYIAQPLDYELQSVNLIKIDARNEPLIAGVYNDSSSTSVIE 604
Qy      557  VTDVNEAPQSOHVFOAKYSEVAIGTKVGNATADPEGLDSSYSLRQDTRGMLKIDHYT 616
Db      605  LVDVDEPPKFEVEGLNANVPENITVGTLLMKAKAPBEKTIKFKNEGHEKMLLNVD 664
Qy      617  GEIFSVAPLDRAGSEYRQVVAATEVYSGSSLSVSEFHLIMDVNDNPRILAKDYGLF 676
Db      665  GEIKTAAALDRVDFHTLTITRYETEGSKMEABMYDHLQVDNNTYRKLQK--TQGR 722
Qy      677  CHPLSAPSLIFEATDDOHLFRGPHFTSLSGSLQNDMEVSKINGTHARLSTRHTDE 736
Db      723  CLQDMTP--LTLTAMDKADPY--GEPTFAISRKS--QNFRIKPVDTGSAKILKKKPS 777
Qy      737  ERAYVVLIRINDGCRPLBGIY--SLPVYTCSCYE--GSCF--RPARGHGTGPTVMANGILL 793
Db      778  EQNVATVPINVLDMAG--LGIQKFPVRIKCNCTKGYCIYPASHSKWL--SMGSTIGILA 833
Qy      794  TLLVYIGIILAVVIRIKKQKQDNVESQAASEVKEPL 830
Db      834  GVFGVIGLFGIGLQIKK--KQKORATAREGTTKAM 867

RESULT 3
ID 08UV07  PRELIMINARY;  PRT;  893 AA.
AC 08UV07;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE N-cadherin.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_Taxid=7955;

```

RN [1]
 SEQUENCE FROM N.A.
 RA Lele Z., Folichet A., Concha M., Rauch G.-J., Geisler R., Rosa F.,
 RT Wilson S.W., Bailly-Cuif L., Hammermidt M.;
 "zebrafish paraculture/n-cadherin is required for neural tube closure
 and correct positioning of neurons within the CNS";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF418565; AAL59160.1; -
 DR Interpro; IPR002126; Cadherin.
 DR Interpro; IPR000233; Cadherin_C_term.
 DR Interpro; IPR000215; Serpin.
 DR Pfam; PF00028; cadherin_5.
 DR Pfam; PF01049; Cadherin_C_term; 1.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 5.
 DR PROSITE; PS00232; CADHERIN_1; UNKNOWN_3.
 DR PROSITE; PS50268; CADHERIN_2; 3.
 DR PROSITE; PS00284; SERPIN; UNKNOWN_1'.
 SQ SEQUENCE 893 AA; 98990 MW; 5E3FF0CDD92B3A6F CRC64;

Query Match 16.4%; Score 709.5; DB 13; Length 893;
 Best Local Similarity 27.8%; Pred. No. 4.5e-38;
 Matches 213; Conservative 115; Mismatches 335; Indels 103; Gaps 21;
 QY 123 RPTFLOSKTEGSRONSRRPKPFLYNADL-----DDPATNGQLYQIVQLP 172
 DB 28 QPQFTNEVYVMTADVITTEQVILKVDVDCGRGSLRFESGDPAD----- 73
 QY 173 MINWVYFOINNTGALISLRGSGELNPAKNSYLVISVKGMGQSE-----NSFSDT 228
 DB 74 -----FRL-DADGVTMAAR--TLQTLDRKGS--LEIKAKENSQEQMWHINFTQPK 121
 QY 229 SVDIIV-----TENIWKAPKPEVWENSTDPHP-----IKITQVRMNDPGA 269
 DB 122 QVPVILFPHNSVLVKGDSDVNRKRDMVLP-PNVVLNENRKQPEELVKIQSXKSNLT 180
 QY 270 QVSLV--DKKLPFRPPSIDQ--EGDIYVTPQLDRBEKDAVYFAVAKDEYKELSYPLEI 326
 DB 181 RYSVTGPAGADQNTGFIIDPIISGLSVKPLDREHLPNHLAAHVDINGNQMENPIDI 240
 QY 327 HYKVKQDINNPRTCPSPVTVFEQVEBERLGNSTGLTADRDSENTANFLNRYIEQTP 386
 DB 241 IINVDINMNNRBEFTHQINNGVTDSEAKPFTVMTVTSQDKDPNTANGELTKIKLSQTP 300
 QY 387 KLPMDGLFIQTAYAGMLQAKOSLKKQDTPQVNLTEVSDK-----PKTLCFQVQINYI 440
 DB 301 ESSSSSMFTINNTGKIITVAAGLDREKYPQYTLITQADMEGNPTVGSNATAVIRLI 360
 QY 441 DINDQIPTEKSDYGNLTLAEDTNIGSTLITQADADEPFTGSSKILYHIKQDSEGR 500
 DB 361 DVNDNAPEFTRETFHGEVEENRVNVIITMLTV--TDKDEPPTFAMNAVYRIISGDPYGRF 418
 QY 501 GVDTPHTNTGVYIIKKPLDFTFAVNSNIVFKANEPVFGVKYNAASFPAKFTLIVTV 560
 DB 419 SLPTDVTNIEGLTVKVPDFENNSFMTLVADNVPPLASGIRHTRQSTAVVSIRVIV 478
 QY 561 NEAPQSQHVPOAKYSEDAVIGTKVGNATKADPEG---DISYSLGDRGMKIDHYTG 617
 DB 479 NSPNNDPYPKQIKLEBGLPQWSMLTTFNHPDRMQOTISYKLYDPAAMLEIDPNNG 538
 QY 618 ELPFAVPLDREA---GSPYVQVAVTEVGSSLSVSEFHLIMDVNDNPRLLADYTG 673
 DB 539 RISTIAVDRESPPYKGNLYNATFNASDNGVPASGTGLQIYTLIDINDNAPRV----- 592
 QY 674 LFFCHPLSA-----PGLSLEFATDDQHLFRGHFTSLSG--GSLQNDMEVSKINT 724
 DB 593 -----FQSEVCEVEREPENAINITAVDGINPNAGP-YAFELPNRPSDIRRMTLIRISD 647
 QY 725 HARLSTRHTDPERAAVVLIRINDGRPPLGIVSLPTVFCSC-VEGSCFPRAGHOTGP 783
 DB 648 HNGTSLKSYLSGIEYELISITDSGNLPMSTYTIKIRIVCQCDHHDGCV---DMERIM 703
 QY 784 TVGMAVGILLTLLVIGIILAVFIRIKDKKDNVESQAQSEVPR 829

Db 704 AAGTGTGATITATILVIVLWFMVMMKRRDKERQAKQLIDP 749
 RESULT 4
 ID P79883 PRELIMINARY; PRT; 922 AA.
 AC P79883;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Maternally expressed neural cadherin Xmn-cadherin.
 GN Xmn-CADHERIN.
 OS Xenopus.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae.
 OX NCBI_TaxID=8353;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE-TAIL BUD;
 RC MEDLINE=96257964; PubMed=8652409;
 RA Tashiro K., Tool O., Nakamura H., Koga C., Ito Y., Hixasa H.,
 RA Shiohawa K.;
 RT "Cloning and expression studies of cDNA for a novel Xenopus cadherin
 (Xmn-cadherin), expressed maternally and later neural-specifically in
 embryogenesis.";
 RT Mech. Dev. 54:161-171 (1996).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 3 CADHERIN DOMAINS.
 DR EMBL; S62457; AAB37685.2; -
 DR HSSP; P15116; INCI.
 DR Interpro; IPR002126; Cadherin.
 DR Interpro; IPR000233; Cadherin_C_term.
 DR Pfam; PF00028; cadherin_5.
 DR Pfam; PF01049; Cadherin_C_term; 1.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 5.
 DR PROSITE; PS00232; CADHERIN_1; 2.
 DR PROSITE; PS50268; CADHERIN_2; 3.
 KW Calcium-binding; Cell adhesion; Glycoprotein.
 SQ SEQUENCE 922 AA; 101173 MW; FEA7A95CBAF1B640 CRC64;
 Query Match 16.2%; Score 699; DB 13; Length 922;
 Best Local Similarity 26.4%; Pred. No. 2.3e-37;
 Matches 224; Conservative 140; Mismatches 334; Indels 152; Gaps 31;
 QY 5 AHHSLCLIMLYLATNGYQSEKFGSLKPMFTSYEGEPPSQITFQ-FRANPAVTEFLT 63
 DB 34 ASVSSAC-----TPGFSADG--YTLVSP--NIMEGQKLKLVKNGCSSGAGCIWE-T 82
 QY 64 GETDNIPIYIEREGILLYRALDRFRSTHNLQVALDANGIIV---EGPPTIEVNDI 119
 DB 83 NNPD--FVYAGDAVY-----TAREVQIPKQAKFTVAANDHTP----- 120
 QY 120 NDNRPFLQSKYEGSRONSRRPKPFLYNATDLDPAFPGQLYVYQIVQLPMINVMY 179
 DB 121 -----EKWEALIQ-----LFVE-----ETPLNQSQYQ----- 142
 QY 180 FOINNTGALISLRGSGOE-----LNPAPKPSNLYVSVKMGQSENSFSDTSDIT 233
 DB 143 -----TG--SQOSEQSEQOSQSTLLPWRQHNGKLRQKXD----- 176
 QY 224 VTENIWKAPKPEVWENSTDPHPKITQVRMN---DEGAQVSL--YDKEKLPFRPPSIDQ 288
 DB 177 -----WVLP-FVNVPENSRGPPQQLVILRSQKRDITRISITGVGADQPMALFNIDP 230
 QY 289 -EGDIYVTPQLDRBEKDAVYFAVAKDEYKELSYPLEIHVKVDINDNPTCPSPVTVF 347

Db 231 IEGRAWVTRPLDEBESSHYLRAHVDINGNKVENPIDLSIYVIMDNDRPEFFSPING 290
 Qy 348 EVOGENRLGNSIGTLTAHDEBENTANSFLNRYIVEOTPLPMDGLFLIOTYAGMLQAK 407
 Db 291 SYDEAKSPGYVMTVAHADADDINTSGIWMYRIMQSPSPSHDMFVHSKTVINTVA 350
 Qy 408 QSLKQDTPQYNLTIVSDK-----FKLTCVQVINVINDINOQIFPEKSDGNLTAE 461
 Db 351 AGIDRRKVOQYTVVIAQTDWEGNLNGLSNTAATVAVDNDVNNPEFTKME--IGEPV 408
 Qy 462 DTNIGSTIILITQATDADEPFGSSKILYHIKGBSEGLGVDPDPTNTGYVILKPELP 521
 Db 409 ENHVDVAVNLVYVDDQPTSMWNAVKLISSDPDSHFITKIDPVTNEGIVVSKPVYD 468
 Qy 522 ETAAVSNIYKAEPEPELVGVKYNASSFAKFLITVDVNEAPQFSOHVQAVSEDAV 561
 Db 469 EMKSVFPLIWMVNOAPLASSGIOMSLQSTAAVTVSVNDVNEAVFPRNEPIKLGESA 528
 Qy 582 GTKVGAVTADPE---GLDISLARGDTRGMLKIDHTGIFSVAPLDRAGSP----- 632
 Db 529 GRLLITPSAVDPHSMQOVLRYSKISDPANMLAINTNGVSTTAVLDRE--SPFYKDL 586
 Qy 633 YRVQVATEVGGSSLSVSEFHLIIMDVNDNPPR-TAKDYTGLFCHPLSAPSLFEAT 691
 Db 587 YQAKFLATDNGNPPASGTGLLIQLIDINDNAPELPKQAQ---ICERPNGG-INITAL 642
 Qy 692 DDDQHLFRGPHFTFSLGS--GSLQNDWEVSKINGTARSLSTRHTDPEERAYVYLIRINDG 749
 Db 643 DVDRKPSADP-FVFEPLPSVPTTIRNMTIHRINSIYARSLQICYLESQWYDVAVITDS 701
 Qy 750 GRPELBIIVSLPTFGSCVSGSCFRPAGHOTGIPVGMANGLITLLVIGITLAV--F 807
 Db 702 GNPPLVNTSITIKKVCPCDNNGDCTTG---AVALAGLGAIIISILICIIILSWLVP 758
 Qy 808 IRIKODKGD 817
 Db 759 VMNMKREKE 768
 RESULT 5
 090275 PRELIMINARY; PRT; 763 AA.
 AC 090275;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Neural-cadherin precursor (N-cadherin).
 OS Brachydanio rerio (Zebrafish) (zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_Taxid=7955;
 RN [1]
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RC TISSUE=EMBRYO;
 RX MEDLINE=95178741; PubMed=7873785;
 RA Bitzur S., Kam Z., Geliger B.;
 RT "Structure and distribution of N-cadherin in developing zebrafish
 RT embryos: morphogenetic effects of ectopic over-expression";
 RT Dev. Dyn. 201;121-136(1994).
 RL Dev. Dyn. 201;121-136(1994).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PRESENTATIONALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN
 CC NEURONAL RECOGNITION MECHANISM.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY)..
 CC ALL DEEP CELLS BUT LATER BECOMES RESTRICTED TO VARIOUS EPITHELIAL
 CC AND NEURONAL TISSUES. FOUND IN DISCRETE AREAS OF CELL-CELL
 CC ADHESION WHEN EMBRYO IS NEARLY COMPLETED. EXPRESSED IN ADULT
 CC BRAIN, EYES AND TRUNK.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION IS FIRST DETECTED IN THE EMBRYO

CC DURING GASTRULATION.
 CC -1- SIMILARITY: BELONGS TO THE CADHERIN FAMILY.
 DR EMBL; X67648; CAA47890.1; -.
 DR HSSP; P15116; INCI.
 DR ZFIN; ZDB-GENE-990415-171; cdh2.
 DR InterPro; IPR002126; Cadherin.
 DR InterPro; IPR002233; Cadherin_C-term.
 DR Pfam; PF00028; cadherin; 5.
 DR Pfam; PF01049; Cadherin_C-term; 1.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 5.
 DR PROSITE; PS00232; CADHERIN 1; 2.
 DR PROSITE; PS0268; CADHERIN 2; 5.
 KW Cell adhesion; Glycoprotein; Phosphorylation; Transmembrane;
 KW Calcium-binding; Repeat; Signal.
 FT SIGNAL 1
 FT PROPEP 35
 FT CHAIN 36
 FT DOMAIN 36
 FT TRANSMEM 602
 FT DOMAIN 623
 FT REPEAT 36
 FT REPEAT 144
 FT REPEAT 259
 FT REPEAT 374
 FT REPEAT 481
 FT DOMAIN 740
 FT CARBOHYD 149
 FT CARBOHYD 201
 FT CARBOHYD 278
 FT CARBOHYD 331
 FT CARBOHYD 449
 FT CARBOHYD 499
 FT CARBOHYD 528
 FT CARBOHYD 569
 SQ SEQUENCE 783 AA; 87120 NM; B57ECBDF41CCIC5 CRC64;
 Query Match 15.7%; Score 677; DB 13; Length 783;
 Best Local Similarity 30.1%; Pred No. 56-36;
 Matches 187; Conservative 98; Mismatches 288; Indels 48; Gaps 15;
 Qy 239 WKAPKEVENENSTDPHP---IKITQVRWMDGQAYSLV-DKEKLPRFPESIDQ-EGDI 292
 Db 37 WVLP-PVNVLENSRKQFPELVKIQSDKXNTLRKSVTSGPADQPTGLFLIDPISGL 95
 Qy 293 YTTQPLDREKDAYVYVAKBYGKPLSYPLEIHKKYKDINDNPTGCSPTVFEVQGN 352
 Db 96 SVTKPLDRHPIPIFHLRAHVDINGNQMENPIDIINVIDMNDNREFTHQINGTVDG 155
 Qy 353 ERLGNSIGTLTAHDEBENTANSFLNRYIVEOTPLPMDGLFLIOTYAGMLQAKOSLKK 412
 Db 156 AKFGTVMVTISQDKDPTANGMLRYKILSQTPESPSSNMFTINNKTKIITVAAGDR 215
 Qy 413 QDTPQYNLTIVSDK-----FKLTCVQVINVINDINOQIFPEKSDGNLTAEPTNIG 466
 Db 216 EKVPQYTLITQATDADEPFGSSKILYHIKGBSEGLGVDPDPTNTGYVILKPELP 521
 Qy 467 STLTITQATDADEPFGSSKILYHIKGBSEGLGVDPDPTNTGYVILKPELP 521
 Db 276 VTMLTV--TDKQEPGPANNAVRIISGPTGFSIPTPVNIEGLVTVVKKVPDFMNS 333
 Qy 527 SNVFAENPEPELVGVKYNASSFAKFLITVDVNEAPQFSOHVQAVSEDAVITKVG 586
 Db 334 FMLTVVADNEVPLASGIHRTOSTAVSIRVIDNNSPNDPPKIKLEGLPQMSMLT 393
 Qy 587 NVTA-KDEGL--DISYSLRGDTRGMLKIDHTGIFSVAPLDRAGSP-----GSPYVQV 638
 Db 394 TFRHDDPDKYMOQTISYGLVDPANLLEIDPNNGRISTAVLDRESPTVKNLNVATEM 453
 Qy 639 ATVGGSSLSVSEFHLIIMDVNDNPPRLAKDYTGLFCHPLSA-----PSLFEAT 691
 Db 454 ASDNGVPASGCTGLIQLIDINDNAPRV-----FPOBAVCERPEPAINITAV 503

QY 692 DDDQHLFRGPHFTSLGSG--GSLQNDMEVSKINGTHARLSRTHRTDPEERAYVVLIRINDG 749
DB 504 DGLINENAGP-YAFELPNRPSDRIRRWTLTRISGDHQLSLKISYLSGIEYELPISITDS 562
QY 750 GRPLLEGIVSLPYTFPGSC-VEGSCFRPAGHOTGIPYGVMAVGLLTLTLVIGIILAVFTR 808
DB 563 GNLPMNTTYLRLKRVQCCDHGDCV---DMERIMAGLGTGAILIILCIITLILVAVLM 618
QY 809 RIKKDKGKDNVESAQASEVXP 829
DB 619 FVMMMKRDKERQAKOLLIDP 639

RESULT 6
O9HAZ9 PRELIMINARY; PRT; 839 AA.
ID O9HAZ9;
AC O9HAZ9;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Desmocollin 3b.
GN DSC3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:20484176; PubMed:11027496;
RA Whitlock N.V., Hunt D.W., Rickman L., Malhi S., Vogazianou A.P.,
RT Dawson L.F., Eady R.A., Buxton R.S., McGrath J.A.;
RT "Genomic organization and amplification of the human desmosomal
RT cadherin genes DSC1 and DSC3, encoding desmocollin types 1 and 3.";
RT Biochem. Biophys. Res. Commun. 276:454-460 (2000).
CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
DR EMBL: AF293359; AAG3427.1; -.
DR HSSP: P15116; INCU.
DR InterPro: IPR002126; Cadherin.
DR Pfam: PF00028; cadherin; 5.
DR PRINTS: PRO0205; CADHERIN.
DR SMART: SM00112; CA, 5.
DR PROSITE: PS00232; CADHERIN_1; 2.
DR PROSITE: PS0268; CADHERIN_2; 5.
DR Calcium-binding; Cell adhesion; Glycoprotein.
SQ SEQUENCE 839 AA; 93497 MW; F21BE2B865A6855 CRC64;

Query Match 15.5%; Score 671.5; DB 4; Length 839;
Best Local Similarity 30.8%; Pred. No. 1.3e-35;
Matches 187; Conservative 106; Mismatches 263; Indels 51; Gaps 22;

QY 239 WKAPKEVENENTDHPKIKITQVRMNDPQAYSL-----YDKKLPKRPFSIDQSG 290
DB 137 W-APICSGQENSLGFPFLFLOQVR-SDAQNTVTVYSISRGVDKPELNLFTIERD-TG 193
QY 291 DIYVTOPLDREEDKAY--VFYAVAKDEYKGFSLSYPLEIHVKVDINDNPCTCSPYTVRE 348
DB 194 NIECTRPVREEDVDLALAYASTADYSADL--PLPLPRVDEDNHVPFEALVYNE 251
QY 349 VQENELGNSIGTLTHDRDEENTANSFLNRYIVEOTPKLPMDGLFIQTAYMLQALQ 408
DB 252 VESSRPGTGVVAVCATDDEPDMTRKLYSILQOTPRSP--GLPSVSPSTGVITTVSH 309
QY 409 SLKQOTPOYNLTIEVSDKDF-----TLGVOQNVINDINDQIPIFEKSDVGNLTAE 461
DB 310 YLDREVVDKSLMKQDMQDGFGLIGTSTCI--ITVDSNNAPTFQONAE--AFVE 365
QY 462 DTNIGSTILLITQATDADEPTGSSKILYHIKDSSEGRGLVDTPHTNTGVYIILKPLDF 521
DB 366 ENAFNVEILRIPIEDDLINTAMWRVNTLTKENEGHFKISTDKKTNNGSVLSVAPLWY 425
QY 522 ETRAASNIYFKAENPEPLVFGV-KYNAASSPAKFTLLVTVDNAPQSGVHPQAKSVDYA 580
DB 426 EENROYNLIEIGVNNAPFARDIPRVATLNRALVTVAHRLDDEGPECTPAQVYRIKENLA 485

QY 581 IGTKVGNTYAKDPE---GLDISYLRGDTRGMLKIDHVTGTEIFSVAPLDRAGSP-----Y 633
DB 486 VCSKINGYKAYPENNNNGNGLRYKKLHDPKGNITTDIEISGSIITSILDRVETPNEXLY 545
QY 634 RQOVVATEYVGGSSLSVSEFFHLLMDVNDPRLAKDYGLFECFPLSAPGSLIPEATDD 693
DB 546 NITVLAIIDKDDSCCTGTLAVN--IEVDNNDPPEIIOEY--VVICPKRMGYTDIL--AVDP 599
QY 694 DQHLFRGPHFTSLGSG--LQNDMEVSKINGTHARLS-TRHTDPEERAYVVLIRINDG 750
DB 600 DEPVHGAQ-PYFSLPSTSPISRLWSLTKVNDPAALSTYQKNAQOE--YTIPIYKD-- 654
QY 751 RPLLEGIVSLPYTFPGSCVVEGSCFRPAGHOTGIPYGVMAV-GILLTLTLVIGIILAVFTR 809
DB 655 RAGQATKILRLNLCCHCTHPQCRAISRSTGYILGKMAILAILLIGALLFSVLLTVCGV 714
QY 810 IKKDKK 816
DB 715 FGATKKG 721

RESULT 7
O8TBX3 PRELIMINARY; PRT; 713 AA.
ID O8TBX3;
AC O8TBX3;
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Cadherin 13, H-cadherin (heart).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-TESTIS;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC028624; AHH28624.1; -.
SQ SEQUENCE 713 AA; 78320 MW; E947F5FA773988FA CRC64;

Query Match 15.1%; Score 652; DB 4; Length 713;
Best Local Similarity 28.9%; Pred. No. 1.9e-34;
Matches 192; Conservative 100; Mismatches 274; Indels 98; Gaps 21;

QY 160 NQOLYQIVITQPMINNNVYFQINKKTAISL---TRSGQLNPKNPSYN-----LVI 211
DB 61 NDKLRYEV-----SSPYFVNSDGGIYALRNITAVGKTLFVHARTPHADMAELVIV 112
QY 212 SYKDMGGGSEN--SFSDTTSV-----DIYVENIKAPKPVEMVNSDTPHPIKITQVRM 264
DB 113 GSKDQGSLODIFKARISPVPRQKRSIVS-----PILIPENQRPFRDVGK-- 162
QY 265 NDPGAQYSLVDEKLPKRPFP-----SIDQ-----GDIYVTOPLDREEDKAYVF 308
DB 163 -----VSDRPREKSRFLTGKGVDEPBGIFRINENTSSVMTRLDREAVAYQL 213
QY 309 YVAVADEYKPLSYPLEIHVKVDINDNPCTCSPYTVFEVQENELGNSIGTLTAHDD 368
DB 214 FEETTVNGKTLGEPLEIVIIDQNDNRPIFREPGYIGHVWEGSPGTGVVNRMTAFDAD 273
QY 369 EENTANSFLNRYIVEOTPKLPMDGLFIQTAYAG-MQLAKSLKQODT---POYNLTIV 424
DB 274 DATATNALLRNIRQOTDPKSPNNMFYIDPEKGDIVTVVSPALLRETLNPKYELILBA 333
QY 425 SDK-----DFKTLCEVQNVINDINDQIPIFEKSDVGNLTAE DTNIGSTILLITQATDADE 479
DB 334 QMAGDGLVGLGTATATNIMDKNDHSKFTYKGE--QATYBGAVG--YIVLVLYEDKDD 390
QY 480 PFTGSSKILYHIKDSSEGRGLVDTPHTNTGVYIILKPLDRETAASNIYFKAENPEPL 539
DB 391 PATGAMRAYVTLINNGPQSGFEIHTNPQTNBGLSVVPLDYEISAFTLLIKVENEDPL 450


```

QY 540 VEGVKNASSFAKFTLIVDVNAPQFOSQHVFOAKVSEDAVIGTKVGNATKADPEGLD-- 597
DB 451 VPDVSTGSPSTATVHTITVDVNGPVEFPDPMMVTRQEDLSVGSVLTVAATPDLSIQH 510
QY 598 -ISYSLRGDRGMKIDVHTEIFSVAPLDREA---GSPRYQVAVATEVGGSSLSVSE 652
DB 511 TIVSYVYKDPAGMLNINPINGVDTTAVLDRESTFVONSYYTALFLAIDSGNPPATGTGT 570
QY 653 FHLIMDVNDNPPRLAKDVTGLFCHPLSAP-----GSLIFEATDDOHLFRGPHFT 704
DB 571 LITLEDVNDNAP-----FIYPTVAEVCDDAKNLSVILIGASDKDLHPNTDP-FK 619
QY 705 FSLGSSGLOND-WEVSKINGTHARLSRHTDFEERAVVILIRINDGSRPPLGIVSLPVT 763
DB 620 FEIHKAQVDPKVKISKINTHALVSLQ-NLNKANYNLPIMVTDGSKPMTNITDLRQV 678
QY 764 FCSC 767
DB 679 VCSC 682

```

RESULT 8

```

ID 08VDK4 PRELIMINARY; PRT; 714 AA.
AC 08VDK4;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Caderlin 13.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021628; AAH21628.1; -.
DR InterPro; IPR002126; Caderlin.
DR Pfam; PF00028; caderlin; 5.
DR PRINTS; PRO0205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN_1; UNKNOWN_3.
DR PROSITE; PS50268; CADHERIN_2; 3.
DR PROSITE; PS50268; CADHERIN_2; 3.
SQ SEQUENCE 714 AA; 78116 MW; CASCA791C46A2FFB CRC64;

```

Query Match 15.1%; Score 652; DB 11; Length 714;
 Best Local Similarity 29.3%; Pred. No. 1.9e-34;
 Matches 189; Conservative 109; Mismatches 287; Indels 60; Gaps 20;

```

QY 160 NGOLYQIVQIQLPMINNMVFOJNNKTGAISL--TEGSOELNPAKPNSTN---LVI 211
DB 61 NEKLHYEV-----SSPHFKVNSDGLVALRNITAVGRTLFVHARTPHADMALVIV 112
QY 212 SVYDMGQSEN--SFSDTTSV-----DIYVENIMWKAPEKPEMVENSTDPPIKITOVW 264
DB 113 GGDIOGSLDPIEFARTSPVPROKRSIVS-----PLILENORQPPRDVGKVD 164
QY 265 ND--PGAQYSLVND--EKLPRFPFSIDOE--GDIYVTOPLDREEDKAVVYAVADKEXKP 319
DB 165 SDREGSKFRITGKGVQODPKGTFRINENTGSVSTRTLDRETIATQVLYETTDASGT 224
QY 320 LSYPLEIHHVKVDINDNPTCPSPVTVFEVQENRILGNSICTLTAHDREENTANSLY 379
DB 225 LEGVPLEVIVIDNDNRPIFRGPGYIGHVEGSPGTGTWRMTAFADDPATDALLRY 284
QY 380 RIVQOTKRLPMDSGLFLQTVAG-MLQAKOSIKKQDT---PQNYLITEVSGK-----DFK 430
DB 285 NIRQOTDPKSPNNFYIDPEKGDIVTVVSPALLDRETLNKKYELLITLDAQMACLDVGLT 344
QY 431 TLCEVQIVNDINDIQIFEKSDYGNLTADTNIGSTILTITQATDADDEPTGSSKILYH 490

```

```

DB 345 GTATATVIVIDDKNDSPKFTKEF--QATVEGAVG-VIVNLTVEDKODPTTGAMRAVYT 401
QY 491 IIKGDSSEGLGVDTDPHTNTGYIILKKPLDFTAAVSNIVFKAENDEPLVFGKYNASGF 550
DB 402 IINGNPGOSFEIHTNQNEGMLSVKPLDYIISAFHILLIKVENDEPLVDPVSTGSPST 461
QY 551 AKFTLIVDVNAPQFOSQHVFOAKVSEDAVIGTKVGNATKADPEGLD--ISYSLRGDR 607
DB 462 ATVHTITVDVNGPVEFPDPMMVTRQENISVGSVLTVAATPDLSIQHQTIRSVYKDBA 521
QY 608 GMLKIDVHTEIFSVAPLDREA---GSPRYQVAVATEVGGSSLSVSEFHLIMDVNDN 663
DB 522 GMLSINPINGVDTTAVLDRESPFVNSYYTALFLAIDSGNPPATGTGTLITLEDINDN 581
QY 664 PRLAKDVTGLFCHPLSAPGSLIFEATDDOHLFRGPHFTSLGSSGLOND-WEVSKIN 722
DB 582 APVITP--TVAEVCDDAKNLSVILIGASDKDLHPNTDP-FKEIHKQVDPKVKISKIN 638
QY 723 GTHARLSRHTDFEERAVVILIRINDGSRPPLGIVSLPVTFCSC 767
DB 639 NTHALVSLQ-NLNKANYNLPIMVTDGSKPMTNITDLRQVQVSC 682

```

RESULT 9

```

ID 07704 PRELIMINARY; PRT; 814 AA.
AC 07704;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Desmocollin type 2 (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98380123; PubMed=9716266;
RA Roberts G.A., Burdett I.D., Pidgeley S.C., King I.A., Magee A.I.,
RA Buxton R.S.;
RT "Antisense expression of a desmocollin gene in MDCK cells alters
RT desmosome plaque assembly but does not affect desmoglein expression.";
RL Bur. J. Cell Biol. 76:192-203(1998).
CC -1- SIMILARITY: CONTAINS 4 CADHERIN DOMAINS.
DR EMBL; A0002299; CAA03509.1; -.
DR HSSP; P09803; ISUH.
DR InterPro; IPR002126; Caderlin.
DR Pfam; PF00028; caderlin; 5.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS50268; CADHERIN_2; 4.
KW Calcium-binding; Cell adhesion; Glycoprotein.
FT NON TER 1
FT CHAIN 101 814
SQ SEQUENCE 814 AA; 91081 MW; 59EF086C69BD3882 CRC64;

```

Query Match 15.1%; Score 651.5; DB 6; Length 814;
 Best Local Similarity 30.2%; Pred. No. 2.5e-34;
 Matches 181; Conservative 110; Mismatches 256; Indels 53; Gaps 20;

```

QY 239 WKAPKEVEMVENSTDPPIKITQVRWNDPGAQYSL-----VDREKLPFPFSIDOE-- 289
DB 102 W-APICSMQENSLGFPFLFLOIQ--SDTAQNTIYVSRIGPVDR--PKNLFYERDT 157
QY 290 GDIYVTOPLDREEDKAVVYAVADKEXKPLSYPLEIHHVKVDINDNPTCPSPVTVFEV 349
DB 158 GNLFCRPPVDREYSESEIATFATPDGYTPELPLPLVIRIBEDNDNVPITFEKTYVFTI 217
QY 350 QENRILGNSICTLTAHDREENTANSLYRIVYEQPKLPMDSGLFLQTVAGMLQAKOS 409
DB 218 SENCRGSTVGQVCAVDKDEPDMTRLKXSIIEQLPAYFT--LFSMHPATGVITTSSSQ 275

```

```

Oy 410 LKKOQTPQVNLITVEBDXO-----FKTLGCVQJNVINDINOJIFKESDVGNLTLMBDN 464
Db 276 LDRELLDKQÜLKIKQVDDMGQYFGLQTTSTCIINIDVNDNLFTFRITST--VTSVEANT 333
Oy 465 IGSILLITQATDADEBPTGSSKILYHIKIGDSEGRGLVDDTPHTNNGYVILIKKPLDPETA 524
Db 334 VDVEILRATVVEDKDLMTANMFRANYITLIGNENENPKIYTDPKTNBSILCVAPLVAPEER 393
Oy 525 AVSNIVFKAEDEPL-VFEGVKTNASSFAKFTLLIVTDVNEAPQFSOHVQAKVSEDAVIGT 563
Db 394 KQVDDQIVVNNAPYKESKESLSTSTWATVTVANNQDDEGPECSPRQÜTQIOTENPVGT 453
Oy 584 KVGNTYADKPE---GLDLSYSLRGDRGMÜLKIDHYVGEIJSVAPLDBREA-----GSPRYQ 636
Db 454 EÜQYKAYDPERRRSSGIRYKÜLSDPKDEKWRINENGSITTFPNLUREAMIRGITYNIT 513
Oy 637 VVATEVGGSSLSVSSEFHLILMDVNNPNPLADYTGFLFCHPLSPAGSLIFPATDDQH 696
Db 514 ILASDKDGRSCYGT--LGITILQDVNDNGSIPKH--TVYICKTVLSBSADIV-ALIDDEP 568
Oy 697 LFRGHPFTFSLGS---GSLQNDMEVSKINGTARLSRHTDPEER-----AYVVLIRIND 748
Db 569 I-NGBPFPDRLENIPDSDIERTRLTKINDTARL-----FYEKDLFPGTYSIRSVAD 621
Oy 749 GGRPPELEGVSLPYPYFCSCY-EGSCRRPAGHQGIGITPYGMAVGLLTLVLIGIILAVVF 807
Db 622 --RHORSILTPKVLVCLDVTENDCLRTARDARGGQDVRIGKWAIIAIIILGIMLFCILF 679

```

RESULT 10			
ID	Q8R490	PRELIMINARY;	PRT; 714 AA.
AC	Q8R490;		
DT	01-JUN-2002 (TREMBlrel. 21, Created)		
DT	01-JUN-2002 (TREMBlrel. 21, Last sequence update)		
DT	01-JUN-2002 (TREMBlrel. 21, Last annotation update)		
DE	T-cadherin.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
SC	TISSUE=HEART;		
RA	Niermann T., Schmutz S.;		
RT	"Cloning of rat T-cadherin.";		
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF494095; AAM14607.1; -		
Q8	SEQUENCE 714 AA; 78085 MW; 7B56142B8162F864 CRC64;		

Query Match	15.0%	Score 650	DB 11	Length 714
Best Local Similarity	29.1%	Pred. No. 2.6e-34		
Matches	188	Conservative 110	Mismatches 287	Indels 60
Gaps	20			
QY	160	NGOLYYIVIQLEMINNNWYFOINNKGTAI---	SLTRSGOELNPAKNPSYV----	LYI 211
Db	61	NKRLHYE-----SSPYFKNSDGLVAVRNTAAGRLLFVHARPHADMAELIV 112		
QY	212	SVKMGGOSEN--SFSDTTSV-----DIIVTENIKAPFVEMVENSNDPHIKITQYRW 264		
Db	113	GKKDIQSSLQDIFKFKARTSPVRQKRSIVS-----FLIPENQSQPFPRDQKVID 164		
QY	265	ND--PGAQYSLVDR--EKLPFPFSSIDOE--GDIYTOPLDRREKQAYFYVAKDEYKGP 319		
Db	165	SDREPGSKFRLLTGKGVDDPKGTFFINETGVSATRLDRRTIATYQLFVETTAASGT 224		
QY	320	LSYPLEIHAKVNDINDNPPTCSPATVEVQENERLNGSIGLTAHREDEENTANSFLNY 379		
Db	225	LEGPAFLFVIVADQNDNRPFRFEGPYIIGHVGEPSGTITVWGMTAFADDDPATIDALLRY 284		
QY	380	RIVGTEPKLPMDGLFLIQTYAG--MQLAKQSLKKQDT--POYNLTIEYSK-----DPK 430		
Db	285	NIRQGTTPKPSNNMYIIDEKGDITVAVSPALDDELTENPKCYLLIEKQADGAGDVGIT 344		

```

Oy 431 TLGVOQINVIDINDOLPIREKSDYGNLTLAEPTNIGSTILQATDDEDEPTGSKLYH 490
Db 345 G7ATVATV1DDKXDSPEKTKKEP---QATVEGAVG-VYVNLVYEDKDDPRTGAMRAYT 401
Oy 491 IINGDEGRHLGVDDTPHTWTGVIIJLKEPDEFTAVASNIIVKAAPEBELVGVXKNASF 550
Db 402 IINGNGGSGFEIHTNPQNTIEGMLSVYKLEIYEAISFHTLLIKENEDPELVDSYGEFSS 461
Oy 551 AKFTLLVTDVNEAPQFSOHVFOAKVSEDAVIGTKVGNVJADPEGLD---ISVSLRGDTR 607
Db 462 ATVH1VLDVNBESPVYPRPMVMTKQENISVGSVLLTVNATDPSLQHOTIRYSVYKDP 521
Oy 608 GMLKIDHTVGEIJSVAPLDREA---GSPYRQVAVATEVGGSSLSVSEFHLIMDVND 663
Db 522 GMLSINPINGVDTTAVLDRBSPFHNVSUYTALFLAIDSGNPPATGCTTLMITLEDVND 581
Oy 664 PPLIADYDGLFFCHPLSAPSLIPEALDDDOHLFRGHFFPSLGSGLOND-WEVAKIN 722
Db 582 APVIYR-TVAVBCVDARNLSVVILIGABDKDILHPETDP-FXFEIHKQIVPDKVWKISKIN 638
Oy 723 GTHARLSTRHDEEBEAYVVLIRINDGSRPPLIEGVSLPVYFEGSC 767
Db 639 NTHALVSLIQ-NINKRANTYMLPMVDSGPEPPTNTITDLRAVYCGC 682

```

Q9WU10	RESULT 11	
ID	Q9WU10	PRELIMINARY; PRT; 4589 AA.
AC	Q9WU10;	
DT	01-NOV-1999	(TREMBLrel. 12, Created)
DT	01-NOV-1999	(TREMBLrel. 12, last sequence update)
DT	01-JUN-2002	(TREMBLrel. 21, last annotation update)
DE	Protocadherin.	
GN	Fat.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostei	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;	
OX	NCBI_TaxID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	SPRAIN-SPRAGUE-DAWLEY;	
RC	MEDLINE=99173791; PubMed=10072790;	
EX	Ponassi M., Jacques T.S., Giani L., French Constant C.;	
RA	"Expression of the rat homolog of the Drosophila fat tumour	
RT	suppressor gene.";	
RL	Mech. Dev. 80:207-212 (1999).	
CC	-1- SIMILARITY: CONTRAST 33 CADHERIN DOMAINS:	
DR	EMBL; AF100960; AAD20459.1; -.	
DR	HSSP; P01132; IEGF.	
DR	InterPro; IPR000152; Asx_hydroxyl.	
DR	InterPro; IPR002126; Cadherin.	
DR	InterPro; IPR000561; EGF-like.	
DR	InterPro; IPR000742; EGF 2.	
DR	InterPro; IPR001881; EGF_Ca.	
DR	InterPro; IPR001791; Laminin_G.	
DR	Pfam; PF00028; cadherin; 33.	
DR	Pfam; PF00008; EGF; 5.	
DR	Pfam; PF00054; laminin_G; 1.	
DR	PRINTS; PR00205; CADHERIN.	
DR	SMART; SMO0112; CA; 31.	
DR	SMART; SMO0179; EGF_CA; 1.	
DR	SMART; SMO0001; EGF_like; 2.	
DR	SMART; SMO0282; LamG; 1.	
DR	PROSITE; PS00010; Asx_HYDROXYL; 1.	
DR	PROSITE; PS00232; CADHERIN_1; 16.	
DR	PROSITE; PS0268; CADHERIN_2; 33.	
DR	PROSITE; PS00022; EGF_1; UNKNOWN_4.	
DR	PROSITE; PS01187; EGF_2; 1.	
DR	PROSITE; PS01187; EGF_CA; 1.	
KW	Calcium-binding; Cell adhesion; EGF-like domain; Glycoprotein	
QO	Hydroxylation; Repeat.	
QO	SEQUENCE 4589 AA; 505987 MW; 8CD0561BDC58677 CRC64;	

Query Match 14.5%; Score 628; DB 11; Length 4589;
 Best Local Similarity 25.3%; Pred. No. 1,3e-31;
 Matches 228; Conservative 135; Mismatches 329; Indels 210; Gaps 32;

QY 26 KFSGLPKMTFSIYE-----GOEPOIIFQ-----FRANPAVTFELTGETDNI FV 72
 DB 2703 RFSEPF--YTYTISEDPITGETIDLRHSGVLTLYLVKNTF-----ESNRDEFV 2754
 QY 73 ERE-GLLYNRALDRETRSTHNLQVAA---LDANGIIVEGPVPTIEVKDINDNRPFLQ 128
 DB 2755 DRQSGRLKLEKSDHETTKWYQFSILARCTLDYEVAAS--IDVSIQVANDNSPVL 2812
 QY 129 SKTEGSRONSRPKPFLVYNATDLDPATNGQLYQVYIQLPMINWYFOINKTKA 188
 DB 2813 NPYEAFIVENLPAGSRVYQVRASDLDSGV--NGQWYSL-DQSQDADIIESFAINMETG 2869
 QY 189 ISLTRREGSOELNAPKPSYMLVSVKDMGQSENSFSDTTSVDIYTE-----NI 238
 DB 2870 ITTLK-----ELDHERASVQIKVVASDHGEKVQ--LSTAIIVDTVTVDNDSPRPTAEI 2923
 QY 239 WKA-----PKVEWENSTDPHPKITQVRNNDPAQVSLVDEKLPFPF-SIDQEG 290
 DB 2924 YKGVSEDDPBGVIALISTDADSEBINR-----QVSYFITGDMALGQFAVENMQNDW 2977
 QY 291 DIYVTOPLDREKDAVYFVAVAKDEYKPLSYPLEIHVKYKDIINDPPTC----- 340
 DB 2978 RYVVKPLDREKDSYLLTYTATD--GTFSKARVEKVLADANDSPVCERTAYSDAIP 3034
 QY 341 ----- 340
 DB 3035 EDALPGKLVMQVATDADIRSNAEITYTLFGSGAEKFKLNDPTGELRTALLDREQAVY 3094
 QY 341 ----- 340
 DB 3095 HLIVKATDGGGRSCQATIVLTLEDVNDNTEPTADPAIATF---ENTERTPTLRQAT 3151
 QY 366 DRDEENANSFLNRIYEQPKLMDGLFLIQTAGMLAKOSLKKODPQVYLTIEVS 425
 DB 3152 DADA--GLNRKISLIDSA-----DQOFSINEQSGILOLEKH-LDRELQAVTYLTILKA 3203
 QY 426 D---KDFKTLCFVQIVNIDINDQIPLEKSDYGNLTAEPTNIGSTILTIOATDADPF 481
 DB 3204 DQGLPRKLTATGTYVSVLIDINDNPVPEEREYEG-AIVSEDIIVIGTEVLQVYAASRD--I 3260
 QY 482 TGSSKILYHIKDGSEGLGVDTDPHNTGYVIIKKPLDEFTAAVSNIVKAEPEPLVF 541
 DB 3261 EANAETIYALISGNEHGFSDID---SKTGALFIENLDYESSHGVLVTEATD----- 3310
 QY 542 GVKVNASSPAKFTLIVTDVNE-APOFSQHVQAVSEDVAICTVAGVNTAKD---PEGLD 597
 DB 3311 GGTPTSDVATVNTIDINDNSPVFSQDSITYVSEDALEQVITIMADADGPNSSH 3370
 QY 598 ISYS-LRGDRGMKIDHVTGEISFVAPLDRAGSPRVQVATEVSGSLSYSEFHLI 656
 DB 3371 ILVSIIEGNGSPPTIDPVREIKVTKPLDEBITISGTLTYQADNDGNPRVNTTYTIN 3430
 QY 657 LMDVNDNPRLAKDYTLFCHPLSAPGSLIFBATDDQHLFRGPHPTFSIGSSLDNDW 716
 DB 3431 VSDVNDNAPLFSSRDNYSVIIQENKPVGFSVLKLVVTKDSSHNQPPFSFALVSGNDNMF 3490
 QY 717 EVSKINGTHARLSTRHDFE--ERAYVVLIRINDGRRPLEGIYSLVTFSCVEGSGCF 774
 DB 3491 EVND---HGVLTLAATVKKRVKOHYLLHVAVADNGKPOLSLTHIDI---RVLESIRP 3543
 QY 775 PA 776
 DB 3544 PA 3545

RESULT 12
 Q9QXA3
 ID Q9QXA3 PRELIMINARY; PRT; 4587 AA.

AC Q9QXA3;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Mouse fat 1 cadherin (Fragment).
 GN MFAT1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=EMBRYO;
 RX MEDLINE=20203461; PubMed=10741417;
 RA Cox B.T.M., Hadjantonakis A.K., Collins J., Magee A.I.;
 RT "Cloning and expression throughout mouse development of mfat1 a
 homolog of the Drosophila tumour suppressor gene fat.";
 RL Dev. Dyn. 217:233-240(2000).
 CC -1- SIMILARITY: CONTAINS 33 CADHERIN DOMAINS.
 DR EMBL; AJ250768; CAB65271.1; -
 DR HSSP; P08709; 1BP9.
 DR InterPro; IPR002126; Cadherin.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF Ca.
 DR InterPro; IPR001791; Laminin G.
 DR Pfam; PF00028; cadherin; 33.
 DR Pfam; PF00008; EGF; 4.
 DR Pfam; PF00054; laminin G; 1.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 31.
 DR SMART; SM00179; EGF CA; 1.
 DR SMART; SM00001; EGF-like; 1.
 DR SMART; SM00282; LamG; 1.
 DR PROSITE; PS00222; CADHERIN_1; 16.
 DR PROSITE; PS00268; CADHERIN_2; 33.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
 DR KEGG; K04102; Calcium-binding; Cell adhesion; Glycoprotein.
 FT NON TER 4587 4587
 SQ SEQUENCE 4587 AA; 506036 MW; 4D3F23B05127C84 CRC64;

Query Match 14.4%; Score 624; DB 11; Length 4587;
 Best Local Similarity 25.1%; Pred. No. 2.5e-31;
 Matches 227; Conservative 135; Mismatches 329; Indels 212; Gaps 33;

QY 26 KFSGLPKMTFSIYE-----GOEPOIIFQ-----FRANPAVTFELTGETDNI FV 72
 DB 2704 RFSEPF--YTYTISEDPITGETIDLRHSGAVLYLVKNTF-----ESNRDEFV 2755
 QY 73 ERE-GLLYNRALDRETRSTHNLQVAA---LDANGIIVEGPVPTIEVKDINDNRPFLQ 128
 DB 2756 DRQNGRLKLEKSDHETTKWYQFSILARCTLDYEVAAS--IDVSIQVANDNSPVL 2813
 QY 129 SKTEGSRONSRPKPFLVYNATDLDPATNGQLYQVYIQLPMINWYFOINKTKA 188
 DB 2814 SPYEAFIVENLPAGSRVYQVRASDLDSGA--NGQWYSL-DQSQDADIIESFAINMETG 2870
 QY 189 ISLTRREGSOELNAPKPSYMLVSVKDMGQSENSFSDTTSVDIYTE-----NI 238
 DB 2871 ITTLK-----ELDHERASVQIKVVASDHGEKVQ--LSTAIIVDTVTVDNDSPRPTAEI 2924
 QY 239 WKA-----PKVEWENSTDPHPKITQVRNNDPAQVSLVDEKLPFPF-SIDQEG 290
 DB 2925 YKGVSEDDPBGVIALISTDADTEINR-----QVSYFITGDMALGQFAVENMQSDW 2978
 QY 291 DIYVTOPLDREKDAVYFVAVAKDEYKPLSYPLEIHVKYKDIINDPPTC----- 340
 DB 2979 RYVVKPLDREKDSYLLTYTATD--GTFSKARVEKVLADANDSPVCERTAYSDTIP 3035
 QY 341 ----- 340
 DB 3036 EDALPGKLVMQVATDADIRSNAEITYTLFGSGAEKFKLNDPTGELRTALLDREQAVY 3095
 QY 341 ----- 340

Db 3096 NLVKAIDGGRCQAIVLTLEDVNDNADEFTAEPTITVF---ENTREGETPLTRVQAT 3152
 QY DRDEENTANSEFLNRIYVEOTPKLPMOGLPLIQVAGMLQAKOSLKQODTPQYNLTIEVS 425
 Db 3153 DAD--TGLNRKISYSLVESA-----DQOFISINERSGITOLEKH-LDELEQAVTTLTKAV 3204
 QY 426 D----KDFKTLCEVOQINVIDINDOIPFEKSDYGNLTAEADTNIGSTILTITQATDADEPF 481
 Db 3205 DQGLPRLTATGTVVSVLDINDNPPVEFREYEG-ASVSEDIYIGTEVLQVNAASRD--1 3261
 QY 482 TGSKLLIHLIKDSEBGLGVDTPHTNTGYIIKKPLDETAASNIYKAEENPEELVF 541
 Db 3262 ENAAETIYAIISGNHKGESID-----SKTGAFIIBSLDYESSHEVYLVEARD----- 3311
 QY 542 GVKYNSAFKFLIYTDVNE-APQSOHVEQAKNSDVNIGTKVGNVATKD---PEGLD 597
 Db 3312 GGTPLSLSDVATVNTVINDNPSVPSQDTYTTVSEEDALEQPVITIMADDADGPNNSH 3371
 QY 598 ISYS-LRCDTRGWLKIDHTGEIFSVAPLDRAGSPRYOVATEVGGSSLSVSEPHLI 656
 Db 3372 IHYSIIIEGNGSPFTIDPVKGEVKTKEPLDRETIISGTLTLVQAADNPNPRVNTTIVNID 3431
 QY 657 LMDVNDNPPRLAKDYTGLEFCPLSAPGSLIFEATDDQHLFRGPHFTSLGSSQLQNDW 716
 Db 3432 VSDVNDNAPLFSHDNYSVLIQENKPVGFSVLKLVTDKDSHNGPPEFFITVSGDENAP 3491
 QY 717 EYKINGTARLSTRHTDPEERA---YVULIRINDGGRPLEGIVSLPTVFCSCVSGSC 773
 Db 3492 EVNQ---HGVLLTAAT-IRKVKDHYFLHVKVDGSKQLSMTHIDI---RVIBESIH 3543
 QY 774 RPA 776
 Db 3544 PPA 3546

RESULT 13
 ID Q8R508 PRELIMINARY; PRT; 4555 AA.
 AC Q8R508;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Fact3.
 GN PAT3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY, TISSUE=BRAIN;
 RX MEDLINE=21670969; PubMed=11811999;
 RA Mitui K., Nakajima D., Ohara O., Nakayama M.;
 RT "Mammalian fact3: a large protein that contains multiple cadherin and
 RT EGF-like motifs.";
 RL Biochem. Biophys. Res. Commun. 290:1260-1266(2002).
 DR EMBL: AB076401; BAB86869.1;
 SQ SEQUENCE 4555 AA; 502077 MW; B47C8C10638EA73B CRC64;

Query Match 14.1%; Score 611; DB 11; Length 4555;
 Best Local Similarity 25.4%; Pred. No. 1.7e-30;
 Matches 225; Conservative 144; Mismatches 320; Indels 196; Gaps 33;
 QY 30 PLKPMFTSIEGQSPQIIFQKAPNPAVTFELTGTDNIPVIERE-GLLYNRALDRET 88
 Db 2733 PMSQVRFSTVNGERPE-----NKNENVFIEGTGTAKIKLDKRLDHEV 2774
 QY 89 RSTNHLQVNA---LDAAGIIVEGPPVITILEYKIDNINRTPLOSKREGSVRQNSRPGKPF 145
 Db 2775 SPAPFKVAATPLDLDVDIVF--TVDDVVKVLDLNDNKVEFTSSYETIIMGMPVGYTL 2832
 QY 146 LVVNAATDLDPATPNQGLYYQIVIVQLPMINNVYFQINNKTAISLTREGSGLNPAKNP 205

Db 2833 AQVRAIDTWMGA--NQGVTSLSHSDHLEKWEAFNIDSWTGMISTLK-----DLDEHTDP 2886
 QY 206 SYNLVSVYDMDGQSGSNSDSTTSVUIITENIKWAKPKPVEMENSTDPPIKITQVRAN 285
 Db 2887 TFSFFVVASDLG--EAFSLSSMALVSVKVIDINDNAP-----VFAHEVVRGNVKES 2935
 QY 266 DPG-----AQSLVDEKEL-----PRPFS--IDQEGDIYVTOPLDREKD 304
 Db 2936 DPGEVVAVALSTLDKOTSINROVSHITGNGRPFALGAMQSEKVKVKEPLDREED 2995
 QY 305 AYVFAVADKEYKPLSYLEIHKVKVDINDNPTC-----PSPTVVEVOE 351
 Db 2996 IYFLNITASD--GLFPTQAM-VEVTVSDVNDNSPVCDOVAYASASLPEDIPSNKILIKVSA 3052
 QY 352 NER-----LGNS-----IGTLAHRDE----- 369
 Db 3053 KADIGSNGDIRYSLVSGNSDFLDPESGELTTLALDRERVPVYNLARATDGGGRFC 3112
 QY 370 -----ENTANSFLNRIYVEOTPKLPMO----- 391
 Db 3113 SSTVLLLEDVNDNPPVSSNHTACVIENTATKALLTRVQAVDPVGINRKVYSLSEDS 3172
 QY 392 --GLPLIQVAGMLQAKOSLKQODTPQYNLTIEVSD---KDFKTLCEVOQINVIDINDO 445
 Db 3173 ASGVFSIDSSGVIVL-EQPLDREOQSSYNISVRATDQSGOGLSLSTVITIVLDINDN 3231
 QY 446 IPFEKSDYGNLTAEADTNIGSTILTITQATDADEPTGSSKLLIHLIKDSEBGLGVDTD 505
 Db 3232 PVFERRDY-LTVFEDTSLGTQVLSVFATSKD--IGTNAETIYLIIRSGNEQCKFPI--- 3285
 QY 506 PHNTGCVIIEKKPLDETAASNIYKAEENPEELVGVKYNSSFAKFLIYTDVNE-AP 564
 Db 3286 -NKTGGSIVLEADLDEMCKRFLVVEAKD-----GTPALSTAATVSIIDLVDVNDNP 3338
 QY 565 QFSOHVFOAKVSEDVNIGTKVGNVATKD---PEGLDISYSL-RCDTRGWLKIDHTGEI 619
 Db 3339 RFSQDVSAVISEDALDEGDSVILLIAEDVDSKPNQ-QIRFSIVGDRDNFAVDPILGLV 3397
 QY 620 FSVAPLDRAGSPRYOVATEVGGSSLSVSEPHLIIMDVNDNPPRLA-KDYTGLEFCFH 678
 Db 3398 KYKKKILDRERVAGYSLILQAVSGIIPAMSGITVINDIDVNDNSVFTPANTAVIQEN 3457
 QY 679 PLSAPGSLIFEATDDQHLFRGPHFTSLGSSQLQNDWVSKINGTHALSTRHTDPEER 738
 Db 3458 KPVGTSLIQLVVTDRS-FHNGPPPSFSLSGNEDEEFWLDHGILRSVVRHMSPE- 3515
 QY 739 AYVULIRINDGGRPLEGIVSLPTVFCSCVSGSCFPAPAGHOTGIP 783
 Db 3516 -YLCITQADSGKPOQ---VSHITIRVRIESTHKP---TAIP 3552

RESULT 14
 ID Q9VSN8 PRELIMINARY; PRT; 3579 AA.
 AC Q9VSN8;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE PROTOCADHERIN-like wing protein STAN precursor (STABRY NIGHT
 DE protein) (FLAMINO protein).
 GN STAN OR CG11895 OR FMT.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC TISSUE=EMERYO;
 RX MEDLINE=20025940; PubMed=10556066;
 RA Chae J.W., Kim M.-J., Goo J.H., Collier S., Gubb D., Charlton J.,
 Adler P.N., Park W.J.,

FT CARBOHYD 1441 1441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1650 1650 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1678 1678 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1747 1747 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1843 1843 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1975 1975 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2016 2016 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2028 2028 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2071 2071 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2088 2088 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2196 2196 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2320 2320 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2784 2784 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3083 3083 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3228 3228 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 181 181 T -> S (IN REF. 1).
 FT CONFLICT 361 361 Q -> L (IN REF. 1).
 FT CONFLICT 395 401 MSLSLS -> NGITVGLP (IN REF. 2).
 FT CONFLICT 1968 1968 Q -> H (IN REF. 1).
 FT CONFLICT 2271 2271 G -> E (IN REF. 1).
 FT CONFLICT 2502 2502 R -> C (IN REF. 2).
 FT CONFLICT 2627 2627 D -> G (IN REF. 2).
 FT CONFLICT 2709 2709 T -> S (IN REF. 2).
 FT CONFLICT 2756 2756 Q -> R (IN REF. 2).
 FT CONFLICT 2901 2901 C -> Y (IN REF. 1).
 FT CONFLICT 3098 3098 L -> P (IN REF. 1).
 FT CONFLICT 3569 3579 ERNIDDEDTTV -> DSEABY (IN REF. 2).
 SQ SEQUENCE 3579 AA; 397139 MW; 48801C493031FB19 CMC64;

Query Match 13.9%; Score 601.5; DB 5; Length 3579;
 Best Local Similarity 29.1%; Pred. No. 5e-30;
 Matches 224; Conservative 131; Mismatches 321; Indels 95; Gaps 35;

QY 37 SIYEGBPQIIQOFKAP-----PAYFELGEMDN-----IFVIE-REGILY 79
 DB 472 SIIEGATVSGTTLTATDODICKNAIEIGIAVTDAGLAODQEMPIFRIDSRGVS 531
 DB 80 YNRALDRSTRNTLQVAALDANGIIVE---GPVITIEKQDINDNRPFLQSKYGSVR 136
 DB 532 TSSSLURETSDSHLVLTADLASQSERATATASQVAVLDNDNDYPPQSEETTYVQVP 591
 QY 137 QNSRPG---KPELYNATDIDDPATNGQLYQIVQLPMINNVYFOJNNKTAISLTR 193
 DB 592 EIQGMGTEDNTVAHIRATDQ---GNNAIRYALIG3---NTQSFISDMSGDVSLVK 645
 QY 194 EGSQELNPAKNBSYMLVISKMGQSGSEFSDTTSVDIIVENIKAKR-----PVEM 247
 DB 646 ---PUDYESVRSLVIRAQDGGSPSR---SMTQLLVAVDANDNAPREFYTSQFQESV 698
 QY 248 VENSTDPHPKILITQVAMNDPGAQ---YSLVNDKEKLPRFPFSID-OEGDIYVTPIDREE 302
 DB 699 LENVPGVMIIRQAQVDSDEGNAELTYSISED--DNFPLAVDPRTGAVQITKPLDREE 756
 QY 303 KQAYFYVAQDEYKGLSYPLEIHVKVDINDNPPTCSPTVFEVQENELGNSIGTL 362
 DB 757 QGFAPQVAVAKGQGVPPKSASSSVITVQDVNDNDNAPFKYEAUVGDEQPGTEVTV 816
 QY 363 TAMDREENTANSFLNRYVEQPKLPMOGLPLI--QYAGMLQALAKSLIKODTQVYL 420
 DB 817 TADPDED---SRLHYETTGNTNR---GRFITSQNGRGLITTA-QSILDVQKQERFLL 867
 QY 421 TIEVSKDEFT-LCFQVQVINDINDQIFEKSDYGLTLAEDPTNGSTLITQIADAE 479
 DB 868 TYAATDSGSRDPAVTHINITDANNAPFEENAPY-SASVEDAPVGTIVLVASIDSD- 925
 QY 480 PPTGSKILYHIKDGSEKGLVDTP---HTNTGVYIIKKPLDEFETAASNIIVKXENP 536
 DB 926 --GVVAQITYSINERSINGLG-SPDPFSINPQTGAIVTNAPLDRITTSGYLLTVAKO- 981
 QY 537 EPLVGVKNASSFAFTLIVDUNE-APQFSGHVQAVSDVAIGTVKGVNTADPE- 594
 DB 982 -----GNPISLDTTDEIGVTDVNDNAPAFKSPLYQASILDALVGTISVIOVAASDPDV 1036

QY 595 GLD--ISYSL--RGDRGMKIDHTVGEIFSVAPLDRREAGSPYRVQVAVTEVGGSSLSYV 650
 DB 1037 GUNGRKYLISDRDIDSGFVIDPISGTRTKNGDRESVAVHLLAIVDKSPPLSLST 1096
 QY 651 SEPHLLIMVNDNPPLAVDTGLFCHPLSAD-GLIFE--ATDDQHLFRGPHFTFSI 707
 DB 1097 VEQIRLEEDVNSPPFPASDKITLYV--PENSFVSGVGEIHADDEGNNAVH--YGI 1152
 QY 708 GSGSLQNDWE-VSKINGTARLST-RHTDPE--ERAYVYLIRINDGCRPL 754
 DB 1153 IGGDSNAPSLVTRPESRRAQLTETLDESTRKFEELVVR--AASPL 1200
 RESULT 15
 ID 09VW71 PRELIMINARY; PRT; 4643 AA.
 AC 09VW71;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DE Putative fat-like cadherin precursor (CG7749 protein).
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCBT_Taxid:7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktiroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockett P., Brothier P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu L.B., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasco P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J.C., Moshrefi A.,
 RA Mount S.W., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclab J.M.,
 RA Palczolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reiert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
 RA Wang Z.-Y., Weisstein D.A., Weislock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.A., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers R.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -!- FUNCTION: MAY BE INVOLVED IN CELL ADHESION.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE CADHERIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.

-1- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAIN.
 CC EMBL; AE003515; AAP49078.1; -.
 DR HSSP; P15116; INCI.
 DR FlyBase; FBgn0036930; fat2.
 DR InterPro; IPR001152; Asx hydroxyl.
 DR InterPro; IPR002126; Cadherin.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF CA.
 DR InterPro; IPR001791; Laminin_G.
 DR Pfam; PF00028; cadherin; 31.
 DR Pfam; PF00008; EGF; 5.
 DR Pfam; PF00054; laminin_G; 1.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 33.
 DR SMART; SM00179; EGF_CA; 1.
 DR SMART; SM00001; EGF-like; 5.
 DR SMART; SM00282; Lamg; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 1.
 DR PROSITE; PS00232; CADHERIN_1; 17.
 DR PROSITE; PS00268; CADHERIN_2; 41.
 DR PROSITE; PS00022; EGF_1; 5.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01187; EGF_CA; 1.
 KW Hypothetical protein; Cell adhesion; Signal; Transmembrane; Repeat;
 KM EGF-like domain; Glycoprotein; Calcium-binding.
 FT SIGNAL 1 35 POTENTIAL.
 FT CHAIN 36 4643 PUTATIVE FAT-LIKE CADHERIN.
 FT TRANSMEM 36 1647 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1648 1668 POTENTIAL.
 FT DOMAIN 1669 1699 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 169 180 CADHERIN 1.
 FT DOMAIN 189 288 CADHERIN 2.
 FT DOMAIN 293 397 CADHERIN 3.
 FT DOMAIN 406 504 CADHERIN 4.
 FT DOMAIN 513 610 CADHERIN 5.
 FT DOMAIN 619 713 CADHERIN 6.
 FT DOMAIN 718 874 CADHERIN 7.
 FT DOMAIN 883 977 CADHERIN 8.
 FT DOMAIN 986 1085 CADHERIN 9.
 FT DOMAIN 1094 1190 CADHERIN 10.
 FT DOMAIN 1199 1296 CADHERIN 11.
 FT DOMAIN 1305 1402 CADHERIN 12.
 FT DOMAIN 1413 1503 CADHERIN 13.
 FT DOMAIN 1512 1609 CADHERIN 14.
 FT DOMAIN 1618 1714 CADHERIN 15.
 FT DOMAIN 1723 1812 CADHERIN 16.
 FT DOMAIN 1821 1929 CADHERIN 17.
 FT DOMAIN 1948 2030 CADHERIN 18.
 FT DOMAIN 2024 2137 CADHERIN 19.
 FT DOMAIN 2146 2232 CADHERIN 20.
 FT DOMAIN 2247 2338 CADHERIN 21.
 FT DOMAIN 2347 2465 CADHERIN 22.
 FT DOMAIN 2474 2567 CADHERIN 23.
 FT DOMAIN 2576 2670 CADHERIN 24.
 FT DOMAIN 2679 2779 CADHERIN 25.
 FT DOMAIN 2788 2876 CADHERIN 26.
 FT DOMAIN 2885 2983 CADHERIN 27.
 FT DOMAIN 2992 3083 CADHERIN 28.
 FT DOMAIN 3092 3185 CADHERIN 29.
 FT DOMAIN 3194 3289 CADHERIN 30.
 FT DOMAIN 3298 3394 CADHERIN 31.
 FT DOMAIN 3403 3499 CADHERIN 32.
 FT DOMAIN 3508 3604 CADHERIN 33.
 FT DOMAIN 3618 3671 CADHERIN 34.
 FT DOMAIN 3679 3879 EGF-LIKE 1.
 FT DOMAIN 3881 3919 EGF-LIKE 2.
 FT DOMAIN 4129 4166 EGF-LIKE 3.
 FT DOMAIN 4243 4279 EGF-LIKE 4.
 FT DOMAIN 4293 4330 EGF-LIKE 5.
 FT DISULFID 3823 3835 POTENTIAL.
 FT DISULFID 3860 3878 POTENTIAL.
 FT DISULFID 3885 3896 POTENTIAL.

FT DISULFID 3890 3907 POTENTIAL.
 FT DISULFID 3909 3918 POTENTIAL.
 FT DISULFID 4133 4144 POTENTIAL.
 FT DISULFID 4138 4154 POTENTIAL.
 FT DISULFID 4156 4165 POTENTIAL.
 FT DISULFID 4172 4183 POTENTIAL.
 FT DISULFID 4177 4193 POTENTIAL.
 FT DISULFID 4195 4204 POTENTIAL.
 FT DISULFID 4247 4258 POTENTIAL.
 FT DISULFID 4252 4267 POTENTIAL.
 FT DISULFID 4269 4278 POTENTIAL.
 FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 779 779 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 843 843 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 923 923 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1106 1106 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1198 1198 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1312 1312 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1439 1439 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1473 1473 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1511 1511 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 4643 AA; 517614 MW; 7EC52BD36ADD4B89 CRC64;
 Query Match 13.7%; Score 591.5; DB 5; Length 4643;
 Best Local Similarity 24.2%; Pred. No. 3, 4e-29;
 Matches 210; Conservative 153; Mismatches 312; Indels 193; Gaps 33;
 QY 33 PMTFSIVEGO-----EPSQIIFQ-FKANPPAVT-----FELTGETDNIPIVI 72
 Db 2764 PVSEIEIETDANIPEFEKSSVLLKIESTPFGVTLKHMIGNYTFKSIADQDH-FMI 2822
 QY 73 EREGLLYNRADRETRSTHNLQVALLDANGIIVEGPVITTEVKDINNRPTELKSYE 132
 Db 2823 SPSEGIILQOTLDREQOESHNLVVAETSTVPFAYADVLLDVRDENDYFEFNTFRS 2882
 QY 133 GSVRNSRPGKPEFLYVNAATLDDEPATNGOLYQVYIQIPMINNWAYFOINNTGAIISLT 192
 Db 2883 ASVAENSEKVISLVKVSATDAD--TGPNGDIRYLESPTENIONI--FDIDISGWITLL 2938
 QY 193 REGSQELNPAKPSYNLVISVKDMGQSENSFSDTTSVDIIVTEN--IKWP-KVEVQVE 249
 Db 2939 TSLDRFV---QSEYFNKYIADNGHPKDKAPVTKIVDYNDAPVFLPIEGSLVEE 2994
 QY 250 NSTDEPRPIKITQVRNMDPGAQYSLVDEKELPR-----PSIDQEGDIYVTOPLDRE 301
 Db 2995 NAL-PGTVLV-LMLLIDPD---LEKQWDFEIVSGDKQAOFOJQKSGELFAKPLDRE 3047
 QY 302 EKDAYVFVAADKDEGKPLSYPLEIHVKVNDINDNPPCTSPSPVTFEVEQENERLGNISGT 361
 Db 3048 QLMFNLISITATD--GR-FLAKANVEIDVYKDINDNPPCYLKPKNHSTNESISIGTTIVE 3104
 QY 362 LTAHDEBENTANSPLNRYIVEQTPKLPMDGLELIQTVAGMLQALAKSLKKQDTPQYNTLT 421
 Db 3105 VKAIDDFQSKFLFVLSGKAD-----FSIGKSGILKVA--SALDRETFPKYKLV 3154
 QY 422 IEVSD-KDRTLCF--VQINVIDINDQIPFEKSDVGNLTLLADTNIGSITLITQATDAD 478
 Db 3155 AHVQDGDVDFQECFSEIITVNDINDNMPIFSAQY-RVSPEDAQLNTLITVHAMDKD 3213
 QY 479 EPTGSSKILYHIKIDSEGRGLGVDTPHNTGTVYIIKKPLDEFAVANSIVFEKAE---- 534
 Db 3214 ---FGVNRQIKYLSMBENHDYKIS---KSTQIILHLSLDEFTISLPLVYAKDECGV 3266
 QY 535 -----NPEPLVEGVK----- 544
 Db 3267 PKLHSIATVAVNLTLDNDNPPERSMROYCKILENATHGTEVCYKVAATSIDGVNADIHY 3326
 QY 545 -----YNASSF-----AFITLIVTD 559
 Db 3327 FIMSGNEGKFKMDSTTGDLVLNATLDYEMSKFVPLTIOAIDGTPPLSNNAVNVNISILD 3386

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: June 22, 2003, 00:55:44 ; Search time 44 Seconds
(without alignments)
1817.815 Million cell updates/sec

Title: US-10-025-380-1081

Perfect score: 4321
Sequence: 1 MILDHLHSLCLIMLYLATG.....DKGKDNYESAQAASEVKPLRS 832

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

```
Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4303	99.6	832	2	S55396	Li-cadherin - human
2	3414	79.0	827	2	A53954	Li-cadherin precursor
3	954	22.1	829	2	I46536	KSP-cadherin - rat
4	715	16.5	913	1	IUCHR	R-cadherin precursor
5	705	16.3	913	1	A47543	R-cadherin precursor
6	699.5	16.2	906	1	IUHUN	N-cadherin 2 precursor
7	696	16.1	912	1	IUCHN	N-cadherin precursor
8	694.5	16.1	906	1	IUMSN	N-cadherin precursor
9	688.5	15.9	877	1	IUBON	N-cadherin precursor
10	684.5	15.8	916	2	C38992	cadherin 4 precursor
11	677	15.7	783	2	I50116	N-cadherin precursor
12	671.5	15.5	826	2	B55353	desmocollin, type
13	671.5	15.5	896	2	A55353	desmocollin, type
14	667.5	15.4	847	1	IUHUB	desmocollin, 3b precursor
15	667.5	15.4	901	1	IUHDA	desmocollin 3a precursor
16	663.5	15.4	906	1	IUXIC2	N-cadherin 2 precursor
17	661	15.3	896	2	I45858	desmocollin - bovin
18	657	15.2	713	2	B38992	cadherin 13 precursor
19	653.5	15.1	809	1	IUBOD	desmocollin 2b precursor
20	653.5	15.1	863	1	IUBOC	desmocollin 2a precursor
21	647.5	15.0	905	1	IUXIC1	N-cadherin 1 precursor
22	634	14.7	887	1	IUCHL	E-cadherin precursor
23	630.5	14.6	712	1	IUMST	T-cadherin precursor
24	630.5	14.6	717	2	I51206	T-cadherin 2 - chicken
25	609	14.1	884	2	S34438	uvomorulin - mouse
26	608	14.1	790	2	G02678	cadherin-14 - human
27	604	14.0	884	1	IUMSE	E-cadherin precursor
28	597.5	13.8	732	1	IUCHB	E-cadherin precursor
29	591	13.7	882	1	IUCHC	cadherin 1 precursor

30	585.5	13.6	839	1	IBODP	desmocolin 1b pre
31	580.5	13.4	895	1	IUXCP	sp-cadherin precur
32	580	13.4	730	1	IUMSCP	M-cadherin - mouse
33	579.5	13.4	871	2	S47518	cadherin - African
34	576	13.3	761	1	IBODE	desmocolin 1a - b
35	574.5	13.3	796	2	A53584	OB-cadherin precur
36	574.5	13.3	840	2	I37281	Dscia precursor -
37	574.5	13.3	894	2	I37282	Dsc1b precursor -
38	572.5	13.2	796	2	I49556	cadherin-11 - mouse
39	571.5	13.2	796	2	I48277	cadherin-11 - mous
40	571	13.2	3034	2	T14119	seven-pass transme
41	570	13.2	770	2	B48910	desmocolin 1b pre
42	570	13.2	824	2	A48910	desmocolin 1a pre
43	565	13.1	814	2	G02878	cadherin-15 precur
44	564	13.1	905	2	S43064	cadherin - African
45	560	13.0	822	1	IUMSCP	P-cadherin precurs

ALIGNMENTS

RESULT 1

LI-cadherin - human
C.Species: Homo sapiens (man)
C.Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 31-Mar-2000
C.Accession: S55396
R.Boettlinger, A.; Krefl, B.; Fieger, C.; Dlouhy, B.; Berndorff, D.; Goessner, R.; Taube
submitted to the EMBL Data Library, December 1994
A.Description: Molecular cloning of human LI-cadherin:evidence for a novel type of cadh
A.Reference number: S55396
A.Accession: S55396
A.Status: Preliminary
A.Molecule type: mRNA
A.Residues: 1-832 <BO>
A.Cross-references: EMBL:X83228; NID:9854174; PIDN:CA58231.1; PID:9854175
C.Superfamily: cadherin, cadherin repeat homology
F:455-566/Domain: cadherin repeat homology <CR3>

Query Match	99.6%	Score 4303	DB 23	Length 832
Best Local Similarity	99.5%	Pred No. 1,3e-263		
Matches 829	Conservative	2	Mismatches 2	Indels 0
			Gaps	0
Qy	1	MILQALHSLICLLMLYLATGYGQEGKFSGPLKPMTFISYIEGQSPQIIIFQKAMPAAVTF	60	
Dp	1	MILQALHSLICLLMLYLATGYGQEGKFSGPLKPMTFISYIEGQSPQIIIFQKAMPAAVTF	60	
Qy	61	ELTGEIDNIFVIERBELLYNPAIDRETRSTHNLQVALDANGIIIVGSPVITTEVKDIN	120	
Dp	61	ELTGEIDNIFVIERBELLYNPAIDRETRSTHNLQVALDANGIIIVGSPVITTEVKDIN	120	
Qy	121	DNRPTFLQSYKESVYKQNSRPGKPFYYNATDLPATPNQQLYYQIYIOLPMNNMYF	180	
Dp	121	DNRPTFLQSYKESVYKQNSRPGKPFYYNATDLPATPNQQLYYQIYIOLPMNNMYF	180	
Qy	181	QINNKTAISLTREGSQELNPAKNPSYNLVLSYKMGQSGSENSFSDTTSVDIIITENIMWK	240	
Dp	181	QINNKTAISLTREGSQELNPAKNPSYNLVLSYKMGQSGSENSFSDTTSVDIIITENIMWK	240	
Qy	241	APRPVEMVENSTDPHPKIKITQVRANDPGAQISLVNDEKLPFPFSPSIDEGDIYVYTOPLDR	300	
Dp	241	APRPVEMVENSTDPHPKIKITQVRANDPGAQISLVNDEKLPFPFSPSIDEGDIYVYTOPLDR	300	
Qy	301	EEKDAVFAVAVKDEGKPLSYPLFIHKVKVDINDNPPTCSPPTVVEVQENBELGNSIG	360	
Dp	301	EEKDAVFAVAVKDEGKPLSYPLFIHKVKVDINDNPPTCSPPTVVEVQENBELGNSIG	360	
Qy	361	TLTAHRDREENTANSEFLANRIVEQRPKLPMDGLFIQTAAQMLQLAQSLKKQDTPQYNL	420	
Dp	361	TLTAHRDREENTANSEFLANRIVEQRPKLPMDGLFIQTAAQMLQLAQSLKKQDTPQYNL	420	
Qy	421	TLVSDICDREKTLCFQVQINVIDINDQIPFPEKSDYGNLTIABDINIGSTILLTQATDAEP	480	

```

Db 421 TIEVSDKDFKLCFVQINVDINDQDPIFEKSDYGNLTIAEDNTIGSTILLTIOATDDEP 480
Qy 481 FTGSSKLLYHIKIKSDSGRLGVDPDPTNTGYVILIKPLDEFTAAVNIVKAEINPELV 540
Db 481 FTGSSKLLYHIKIKSDSGRLGVDPDPTNTGYVILIKPLDEFTAAVNIVKAEINPELV 540
Qy 541 FGVCYNASSPAKFLIYTDVNEAPQSQHVFOAKVSEDVAIGTVGNVTAADPEGLDISY 600
Db 541 FGVCYNASSPAKFLIYTDVNEAPQSQHVFOAKVSEDVAIGTVGNVTAADPEGLDISY 600
Qy 601 SLRGDTGKWLKIDHVTGEISVAPLDREAGSPYRVQVATEVGSSLSVSEFHLIMDV 660
Db 601 SLRGDTGKWLKIDHVTGEISVAPLDREAGSPYRVQVATEVGSSLSVSEFHLIMDV 660
Qy 661 NDNPRLAKDYTGIFCHPLSAPGSLIPEATDDOHLFRGPHFTPSLGSGSLQNDWEVSK 720
Db 661 NDNPRLAKDYTGIFCHPLSAPGSLIPEATDDOHLFRGPHFTPSLGSGSLQNDWEVSK 720
Qy 721 NGTHARLSTRHTFEBERAYVYLIRINDGGRPLEGIVSLPVTFCSCYBGSCEFRPAGHOT 780
Db 721 NGTHARLSTRHTFEBERAYVYLIRINDGGRPLEGIVSLPVTFCSCYBGSCEFRPAGHOT 780
Qy 781 GIPTVMAVGILLTTLVIGIILAVFIRIKDKGKDNVESQAQSEVYKPLRS 832
Db 781 GIPTVMAVGILLTTLVIGIILAVFIRIKDKGKDNVESQAQSEVYKPLRS 832

```

RESULT 2

```

A:Species: Rattus norvegicus (Norway rat)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 31-Mar-2000
C:Accession: A53954
R:Berndorff, D.; Gessner, R.; Kreft, B.; Schroy, N.; Lajous-Petter, A.M.; Reut,
J. Cell Biol. 125, 1353-1369, 1994
A:Title: Liver-intestine cadherin: molecular cloning and characterization of a novel Ca
A:Reference number: A53954; MUID:94266966; PMID:8207063
A:Accession: A53954
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-827 <BBR>
A:Cross-references: GB:X78997; NID:9505562; PIDN:CAAS5631.1; PID:9505563
C:Superfamily: cadherin; cadherin repeat homology
C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; intestine; liver;
F/1-28/Domain: signal sequence #status predicted <SIG>
F/23-827/Product: Li-cadherin #status predicted <MAT>
F/454-565/Domain: cadherin repeat homology <CR3>

```

Query Match 79.0%; Score 3414; DB 2; Length 827;

Best Local Similarity 78.8%; Pred. No. 1e-206;

Matches 655; Conservative 70; Mismatches 102; Indels 4; Gaps 1;

```

Qy 2 ILQAHLSLCLMLYLATGYGQEGKFSGLPKMTFSIYEGQEPQIIFQKAPNAVTFE 61
Db 1 MWSAQHPLFLCLTLTYLGAVGQEGKFSGLPKMTFSIYEGQEPQIIFQKAPNAVTFE 60
Qy 62 LTGTTDNIPIYERGLLYNRALDRETRSTNLTQVALDANGIIVGEGVPITIEVKIND 121
Db 61 LTGTTDNIPIYERGLLYNRALDRETRSTNLTQVALDANGIIVGEGVPITIEVKIND 120
Qy 122 NRPFLLQSKYEGSVRQNSRPGKFLYVNAATLDLPATPNQOLYYQIYIOLPMINNWYFQ 181
Db 121 NRPFLLQSKYEGSVRQNSRPGKFLYVNAATLDLPATPNQOLYYQIYIOLPMINNWYFQ 180
Qy 182 INNKTAISLTRREGSQELNPAKNPSYNLVSVKMGQSGSENSFSDTTSVDIITENIWK 241
Db 181 INNKTAISLTRREGSQELNPAKNPSYNLVSVKMGQSGSENSFSDTTSVDIITENIWK 240
Qy 242 PKPVEAMVENSNDPHPIKITQVRANDPGAQYSLVDKEKLPKRPFSIDQSGDIYVQPLDRE 301
Db 241 PKPVEAMVENSNDPHPIKITQVRANDPGAQYSLVDKEKLPKRPFSIDQSGDIYVQPLDRE 300

```

```

Qy 302 EKDAYVFAVAKDEYKPSLYPLEIHVKYKDINDNPPCPSPVTVFEVENERLGNISIGT 361
Db 301 EKDAYVFAVAKDEYKPSLYPLEIHVKYKDINDNPPCPSPVTVFEVENERLGNISIGT 360
Qy 362 LTAHDEBENTANSFLNRYIVBQTPKLPMDGLFLICTYAGMLQAKSLKKQOTPOYNLT 421
Db 361 LTAHDEBENTANSFLNRYIVBQTPKLPMDGLFLICTYAGMLQAKSLKKQOTPOYNLT 420
Qy 422 IEVSDKDFKLCFVQINVDINDQDPIFEKSDYGNLTIAEDNTIGSTILLTIOATDDEP 481
Db 421 IEVSDKDFKLCFVQINVDINDQDPIFEKSDYGNLTIAEDNTIGSTILLTIOATDDEP 480
Qy 482 TGSSKLLYHIKIKSDSGRLGVDPDPTNTGYVILIKPLDEFTAAVNIVKAEINPELV 541
Db 481 TGSSKLLYHIKIKSDSGRLGVDPDPTNTGYVILIKPLDEFTAAVNIVKAEINPELV 540
Qy 542 GVCYNASSPAKFLIYTDVNEAPQSQHVFOAKVSEDVAIGTVGNVTAADPEGLDISY 601
Db 541 GVCYNASSPAKFLIYTDVNEAPQSQHVFOAKVSEDVAIGTVGNVTAADPEGLDISY 600
Qy 602 LRGDTEGKWLKIDHVTGEISVAPLDREAGSPYRVQVATEVGSSLSVSEFHLIMDV 661
Db 601 LRGDTEGKWLKIDHVTGEISVAPLDREAGSPYRVQVATEVGSSLSVSEFHLIMDV 660
Qy 662 DNBPRLAKDYTGIFCHPLSAPGSLIPEATDDOHLFRGPHFTPSLGSGSLQNDWEVSK 721
Db 661 DNBPRLAKDYTGIFCHPLSAPGSLIPEATDDOHLFRGPHFTPSLGSGSLQNDWEVSK 720
Qy 722 NGTHARLSTRHTFEBERAYVYLIRINDGGRPLEGIVSLPVTFCSCYBGSCEFRPAGHOT 781
Db 721 NGTHARLSTRHTFEBERAYVYLIRINDGGRPLEGIVSLPVTFCSCYBGSCEFRPAGHOT 780
Qy 782 IPTVMAVGILLTTLVIGIILAVFIRIKDKGKDNVESQAQSEVYKPLRS 832
Db 781 IPTVMAVGILLTTLVIGIILAVFIRIKDKGKDNVESQAQSEVYKPLRS 827

```

RESULT 3

```

A:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 05-Nov-1999
C:Accession: I46536
R:Thomson, R.B.; Iggarashi, P.; Blemesderfer, D.; Kim, R.; Abu-Alfa, A.; Soleimani, M.;
J. Biol. Chem. 270, 17594-17601, 1995
A:Title: Isolation and cDNA cloning of Ksp-cadherin, a novel kidney-specific member of
A:Reference number: I46536; MUID:95340560; PMID:7615566
A:Accession: I46536
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-829 <THO>
A:Cross-references: EMBL:U28945; NID:9902885; PIDN:AAC48472.1; PID:9902886

```

Query Match 22.1%; Score 954; DB 2; Length 829;

Best Local Similarity 31.6%; Pred. No. 5.7e-52;

Matches 251; Conservative 142; Mismatches 354; Indels 48; Gaps 22;

```

Qy 62 LTGET-----DNIVIERE-GLLYNRALDRETRSTNLTQVALDANGIIVGEGVPITIEV 116
Db 57 LSGDSGVAAGPESVBAEGFLVTRALDREBOAEQIOVTLAEBOGHVLMGQSVTVHV 116
Qy 117 KOINDRPFPLQSKYEGSVRQNSRPGKFLYVNAATLDLPATPNQOLYYQIYIOLPMINN 176
Db 117 KOINDRPFPLQSKYEGSVRQNSRPGKFLYVNAATLDLPATPNQOLYYQIYIOLPMINN 176
Qy 177 VWFQINNKTAISLTRREGSQELNPAKNPSYNLVSVKMGQSGSENSFSDTTSVDIITVE 236
Db 177 VWFQINNKTAISLTRREGSQELNPAKNPSYNLVSVKMGQSGSENSFSDTTSVDIITVE 235
Qy 237 NIKAKPVEAMVENSNDPHPIKITQVRANDPGAQYSLVDKEKLPKRPFSIDQSGDIYVQ 296
Db 236 STWTPLEPVALLENLKVYPHHLAQVHWSGQDVHYRL--ESQPGPFDVDTGKLYVTG 292

```


J. Cell Sci. 106, 401-409, 1993
 A>Title: Cell binding specificity of mouse R-cadherin and chromosomal mapping of the gene
 A/Reference number: 155501; MUID:94055672; PMID:8270638
 A/Accession: 155501
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-913 <RES>
 A/Cross-references: GB:D14888; NID:9457658; PIDN:BA03605.1; PID:9457659
 C/Comment: Cadherins mediate calcium-dependent intercellular adhesion and are thought to
 C/Superfamily: cadherin; cadherin repeat homology
 C/Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pro
 F:1-27/Domain: signal sequence #status predicted <PRO>
 F:167-166/Domain: propeptide #status predicted <PRO>
 F:167-913/Product: R-cadherin #status predicted <MAT>
 F:167-721/Domain: extracellular #status predicted <EXT>
 F:169-774/Domain: cadherin repeat homology <CR1>
 F:244-249/Region: cadherin binding #status predicted
 F:277-389/Domain: cadherin repeat homology <CR2>
 F:300-304/Domain: calcium binding #status predicted <CR3>
 F:392-504/Domain: cadherin repeat homology <CR3>
 F:507-612/Domain: cadherin repeat homology <CR4>
 F:613-721/Domain: cadherin repeat homology <CR5>
 F:722-753/Domain: transmembrane #status predicted <TM>
 F:754-913/Domain: intracellular #status predicted <INT>
 F:870-885/Region: serine-rich
 F:280,409,554,629,658,659/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 16.3%; Score 705; DB 1; Length 913;
 Best Local Similarity 30.1%; Pred. No. 3,1e-36;
 Matches 178; Conservative 101; Mismatches 280; Indels 33; Gaps 12;

239 WKAPKPEVENSSTDPHPKITQVRW---NDPGAGVSL--VDEKILPRPFPSID-QSGDI 222
 168 WVIP-PIVVENSGRPPQQLVIRSDKMDIPRISITGADQPMPEVFNIDNSGRM 226
 293 YVTOPLREKDAVYFAVAKDEYKELSYLEHVAVKQINDNPPCPSPVIVFEVOEN 352
 227 YVIRPDRERASVHLRAHVNQNGKNVEMPIDIYIVIDNDPRFPIVQVNGSDEG 286
 353 ERLGNSIGTLTAHDREENTANSEFLNRYVEQTPKLPMDGLFIQYAGMLQAKOSLKK 412
 287 SKEGTYMTVTADADSTTANGMVRIRIYVQFQPSQSMFTINSRTGDIIVVAAGLDR 346
 413 QDPFQVNLTEVSGDKD-----FKTLQFQVINVINDIDQPIPEKSPYXGLTLAEDNIG 466
 347 EKVOQYTVIVQATDMEGNLYGLSNVTAITITVDVNDNPEFTSTFAG--EVEPRRIE 404
 467 STILTIQATDADPEPTGSSKILYHIIKGDSEGRAGVTDPHNTGYVILIKKPLDFETAAV 526
 405 TVVANLTVMRDQPHSNMNAVYRIISGDSGHSVATDVTNEGAVTVKAVDEYELNRA 464
 527 SNIVFAENDEPLVFGVKYNASSFAKFTLITVDVNEAPQFSQHVQAKVEDVAIGTKVG 586
 465 FMLTVWVNSQAPLASSGIOMSFQSTAGVTITVTDVNEAPRPPSNKILRIEEGVAGTALT 524
 587 NVTAKDPEGU---DISLSAGDTRGWLKIDHVTGEISVAPLDREA---GSPRYQVVA 639
 525 TFSAVPDRPMQAVRYSKLSDPANMWHINTNSQITTAAILDRESLYTKNNVYEATFLA 584
 640 TEVGSGLSVSEFHLILMDVNDPRLADYGLFCFHPLSAPG--SLIFEATDDOHL 697
 585 ADNGIPRPSGTGLTIQILIDINDAPQLKREAOI-----CERGLNAINITADADMDP 639
 698 FRGPH-FTFSLGSGSLONDEVSKINGTHARLSTRHTFEERAVVILIRINDGRRPLEG 756
 640 NISPVYELPFIPTVAKMWTITRLNGVYAOQLRIILYLEAGYDVAVIITDSGNPLSN 699
 757 IVSLPVTFCSCVSGSCRRPAGHOGIFPTVGMAVGIILTLTLLVIGITLAVPI 808
 700 TSVIKVAVKCPDENGDCITVG--AVAAAGTGGAIVAILIIVITLIMVLL 748

IJHUCN
 Cadherin 2 precursor - human
 N/Alternate names: N-cadherin; neuronal cadherin
 C/Species: Homo sapiens (man)
 C/Date: 30-Jun-1993 #sequence, revision 30-Jun-1993 #text_change 22-Jun-1999
 C/Accession: A38870; S11487; J00751; S13799
 R/Reid, R.A.
 Submitted to the EMBL Data Library, November 1990
 A/Reference number: A38870
 A/Accession: A38870
 A/Molecule type: mRNA
 A/Residues: 1-906 <RE1>
 A/Cross-references: EMBL:X54315; NID:g34998; PIDN:CAA38213.1; PID:g34999
 R/Reid, R.A.; Hempeley, J.J.
 Nucleic Acids Res. 18, 5896, 1990
 A>Title: Human N-cadherin: nucleotide and deduced amino acid sequence.
 A/Reference number: S11487; MUID:91016946; PMID:2216790
 A/Accession: S11487
 A/Molecule type: mRNA
 A/Residues: 1-340 'N', 342-698, 'R', 700-704, 'F', 706-906 <RE2>
 A/Cross-references: EMBL:X54315
 A/Note: this sequence has been revised in reference A38870
 R/Walsh, F.S.; Barton, C.H.; Putt, W.; Moore, S.E.; Kelsell, D.; Spurr, N.; Goodfellow,
 J. Neurochem. 55, 805-812, 1990
 A>Title: N-cadherin gene maps to human chromosome 18 and is not linked to the E-cadherin
 A/Reference number: J00751; MUID:90347462; PMID:2384753
 A/Accession: J00751
 A/Molecule type: mRNA
 A/Residues: 160-194, 'IR', 197-211, 'L', 213-227, 'Q', 229, 'N', 231-235, 'G', 237-248, 'T', 250-356
 A/Cross-references: GB:M34064
 C/Comment: Cadherins mediate calcium-dependent intercellular adhesion and are thought to
 C/Genetics:
 A/Gene: GDB:CDH2; NCAD
 A/Cross-references: GDB:128185; OMIM:114020
 A/Map position: 18q12.1-18q12.1
 C/Superfamily: cadherin; cadherin repeat homology
 C/Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pr
 F:1-27/Domain: signal sequence #status predicted <PRO>
 F:78-159/Domain: propeptide #status predicted <PRO>
 F:160-906/Product: N-cadherin #status predicted <MAT>
 F:162-767/Domain: extracellular #status predicted <EXT>
 F:237-242/Region: cadherin binding #status predicted
 F:270-382/Domain: cadherin repeat homology <CR2>
 F:385-497/Domain: cadherin repeat homology <CR3>
 F:500-605/Domain: cadherin repeat homology <CR4>
 F:606-712/Domain: cadherin repeat homology <CR5>
 F:715-746/Domain: transmembrane #status predicted <TM>
 F:747-906/Domain: intracellular #status predicted <INT>
 F:865-878/Region: serine-rich
 F:190,273,325,402,572,622,651,692/Binding site: carbohydrate (Asn) (covalent) #status p

Query Match 16.2%; Score 699.5; DB 1; Length 906;
 Best Local Similarity 27.4%; Pred. No. 6,7e-36;
 Matches 210; Conservative 122; Mismatches 331; Indels 103; Gaps 21;

117 KDINDNPPFTCSPTVFEVQENRIGNSIGTLTAHDREENTANSEFLNRYVEQTPKLP 389
 47 KDVEGQP-LNVKRF-----SNCKGRKRVQYSSSEPADKRVDEGMYV--AVRSFPLSSB 98
 177 VWFQINNTGAIISLREGSOELNPAKNPSYLVISVKMGQSGSENSPDDTSVDITIV-- 234
 99 HAKFLIY-----AQDETOEKQVAVKLSKLP--LTESVYESAVERIVFP 144
 235 -----TENIWKAPKPEVENSSTDPHPKITQVRW-----DPGAQ 270
 145 RQFSKSHGLQKQKRWVIP-PINLPENSGRPPQQLVIRSDRKNLSLRYVTGPGA- 202
 271 YSLVDEKELPRPFPSIDQ-EGDIYVTOPLREKDAVYFAVAKDEYKELSYLEHVAVK 329
 203 -----DQPTGTFIIPNISGQLSVTKPLDRQIOARFHLRAHVNQNGKNVEMPIDI 256
 330 VKDINDNPPFTCSPTVFEVQENRIGNSIGTLTAHDREENTANSEFLNRYVEQTPKLP 389

A:Residues: 839-906 <MI2>
 A:Cross-references: GB:545011; NID:g256010; PIDN:AAAB23156.1; PID:g256011
 A:Note: Sequence extracted from NCBI backbone (NCBIN:113759, NCBI:P113760)
 C:Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to
 C:Superfamily: cadherin; cadherin repeat homology
 C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pro
 F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-159/Domain: propeptide #status predicted <PRO>
 F:160-906/Product: N-cadherin #status predicted <MAT>
 F:160-714/Domain: extracellular #status predicted <EXT>
 F:162-267/Domain: cadherin repeat homology <CR1>
 F:237-242/Region: cadherin binding #status predicted
 F:237-382/Domain: cadherin repeat homology <CR2>
 F:385-497/Domain: cadherin repeat homology <CR3>
 F:500-605/Domain: cadherin repeat homology <CR4>
 F:606-714/Domain: cadherin repeat homology <CR5>
 F:715-746/Domain: transmembrane #status predicted <TM>
 F:747-906/Domain: intracellular #status predicted <INT>
 F:865-878/Region: serine-rich
 F:130,273,325,402,572,651,692/Binding site: carbohydrate (asn) (covalent) #status predic

Query Match 16.1%; Score 694.5; DB 1; Length 906;
 Best Local Similarity 27.6%; Pred. No. 1.4e-35;
 Matches 211; Conservative 122; Mismatches 332; Indels 99; Gaps 23;

117 KDINDNPTFLQSKYEGSVQRNRPKPFLLVYNAITDLPATPGOLYQIVQLPIMN 176
 47 KDVEGQP-LINVFYSNCRK-----RKQYESSPPADFKVDEGTYV-AVRSPFLTAE 98
 177 VMVQI---NNKTG---AISTREGSQELNPAKNP-SYNIIVISYKDM---GGQSENSE 224
 99 QAKFLIYAQDKETQEKQVAVNLSREPLTREPKEHEIEIYFPPQLAGSALRQK 158
 225 SDTTSVDIIVENIWKAPYEMVENSSTDPPIKITQVRN-----DPGAQYS 272
 159 RD-----WVIP-PINLPENSGPPOELIVRISDRDKNLSRYSVTGPGA--- 202
 273 LVDKEKLPRFPFSIDQ-EGDIYVTOPLDREKDAVVFYAVAKDEYGRKLSPLFHHVK 331
 203 ---DQPTGIFITNPISGQSVTKPLDRELIAFHRLAAHVAVDINGOVENPIDIVNI 258
 332 DINNPPTCPSPVTVFEVQENRIGNSIGTLTAHREDEANTSNFLNRIYEOTPKLPM 391
 259 DMNNRPEFLHQVWNGSVPESSKPTVYMTALTADDPNMLNGLRRLISQAPSPSP 318
 392 GLFLIQTAYAGMLQAKOSLKKQDTPQYNLTIEVSKD-----FKTLCFVOINVIDNQ 445
 319 NMFTINNETGDIITVAAGLDREKVOQYTLITQATDMENPTPYGLSNTATAVITVDVND 378
 446 IPIF-EKSDYGNLTIAEDTNGSTLTITQATDADPEFTGSSKILYHIHKGSEGLGVT 504
 379 PPEFTAMFYESEV--PENRVDVIYANLTITDKQDPHTPANNAIYRISGGDPTRFALIT 435
 505 DPHNTGTVIITKPLDEFTAAVSNIVFAENPEPELVFGVKNASFAKFTLLIVDVNEAP 564
 436 DPNSNDGLVTVYKPIDFTNMFVLTVAAENQVPLAKIGHPPOSTAVSVTVDIVNEP 495
 565 QFSQHVFOAKVSEDAVIGTGVNTAKDPEGL--DISYSLRGDTRGWLKIDHVTGEIS 621
 496 YFAPNPKIIRQEBGLHAGTMTLTLLAODPRYMOQNIIRYTKLSDPANMLKIDPVNGQIT 555
 622 VAPLDREA---GSPYRVQVATEVGVSSLSVSEFHLIMDVNDNPRRLAKDTGLGJFC 677
 556 IAVLDRESPPVQNNINATFTLASDNGIPPMGSGTGLTQIYLDINDNNAQV----- 606
 678 HPLSA-----PGSLIFEATDDOHLFRGPH-FTFSLSGSLQDMVEVSKINGTHALS 729
 607 -PQAEATCETPEPNSINIAADYDIDPNAQBPAPPLPSPTIKKRWITINLNGEPAQIN 665
 730 TRHTDFEERAYVILIRINDGAPPLIGIVSLPVTFQSC-VEGSGCFRPAHQGTGIPV--- 785
 666 LKIKFLKGLIYEVPIITIDSGNPPKSNISIRLVKVCQDSNGDC-----TDVRIYGA 718

QY 786 GMAVGILITLTVIGIILAVFIRIKDKGKDNVESQAQSEVPE 829
 DB 719 GLGTGALIALILCTIITLILVLMFVVMKRRDREDAKQULLDP 762

RESULT 9 IJBON

N-cadherin precursor - bovine (fragment)
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 30-Jun-1993 #sequence_revstion 30-Jun-1993 #text_change 22-Jun-1999
 C:Accession: S11693

R:Law, C.W.; Cannon, C.; Power, M.D.; Kiboneka, P.K.; Rubin, L.L.
 EMBL J. 9, 2701-2708, 1990

A:Title: Identification and cloning of two species of cadherins in bovine endothelial c
 A:Reference number: S11693; MIMD:90360979; PMID:2290969

A:Accession: S11693
 A:Molecule type: mRNA

A:Residues: 1-877 <LIA>

A:Cross-references: EMBL:X53615; NID:g164; PIDN:CAA37677.1; PID:g664894

C:Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought
 C:Superfamily: cadherin; cadherin repeat homology

C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pr
 F:1-130/Domain: propeptide (fragment) #status predicted <MAT>
 F:131-877/Product: N-cadherin #status predicted <EXT>
 F:131-685/Domain: extracellular #status predicted <EXT>
 F:133-238/Domain: cadherin repeat homology <CR1>
 F:208-313/Region: cadherin binding #status predicted
 F:241-353/Domain: cadherin repeat homology <CR2>
 F:356-468/Domain: cadherin repeat homology <CR3>
 F:471-576/Domain: cadherin repeat homology <CR4>
 F:577-685/Domain: cadherin repeat homology <CR5>
 F:686-717/Domain: transmembrane #status predicted <TM>
 F:718-877/Domain: intracellular #status predicted <INT>
 F:836-849/Region: serine-rich
 F:161,244,296,373,543,593,622,663/Binding site: carbohydrate (asn) (covalent) #status p

Query Match 15.9%; Score 688.5; DB 1; Length 877;
 Best Local Similarity 26.7%; Pred. No. 3.1e-35;
 Matches 203; Conservative 126; Mismatches 343; Indels 87; Gaps 19;

126 FLQSKYEGSVQRNRPKPFLLVYNAITDLPATPGOLYQIVQLP-----MINN 176
 7 FPEVDVSAVLSRDLVSGQPLNKFNSC-----NGRKQYESSPPADFKVDEGKVYA 60
 177 VMVQIINNKTGALSTRREGSQELNPAKNPEYNIIVISYKDMGGQSENSEPTSDIIV-- 234
 61 VRSPUSSEHSKFLIYAQDK-----TQEKQVAVKLSKLPALPEBSVKSREIEIYF 115
 235 -----TENIWKAPYEMVENSSTDPPIKITQVRN-----DEGAQ 270
 116 RQYTKNGVYQKQKQKQVWVIP-PINLPENSGPPOELIVRISDRDKNLSRYSVTGPGA- 173
 271 YSLVDKEKLPRFPFSIDQ-EGDIYVTOPLDREKDAVVFYAVAKDEYGRKLSPLFHHVK 329
 174 ---DQPTGIFITNPISGQSVTKPLDRELIAFHRLAAHVAVDINGOVENPIDIVIN 227
 330 VKDINDNPTCPSPVTVFEVQENRIGNSIGTLTAHREDEANTSNFLNRIYEOTPKP 389
 228 VIDMNDNRPPEFLHQVWNGVPESSKPTVYMTALTADDPNMLNGLRRLISQABSTP 287
 390 MDGLFLIQTAYAGMLQAKOSLKKQDTPQYNLTIEVSKD-----FKTLCFVOINVIDN 443
 288 SPNNFTINNETGDIITVAAGLDREKVOQYTLITQATDMENPTPYGLSNTATAVITVDV 347
 444 DQIPIF-EKSDYGNLTIAEDTNGSTLTITQATDADPEFTGSSKILYHIHKGSEGLGV 502
 348 DNPEFTAMFYESEV--PENRVDVIYANLTITDKQDPHTPANNAIYRISGGDPTRFAL 404
 503 DTPHNTGTVIITKPLDEFTAAVSNIVFAENPEPELVFGVKNASFAKFTLLIVDVNE 562
 405 QTDPSNDGLVTVYKPIDFTNMFVLTVAAENQVPLAKIGHPPOSTAVSVTVDIVNE 464
 563 APQSGHVFOAKVSEDAVIGTGVNTAKDPEGL--DISYSLRGDTRGWLKIDHVTGEI 619

```

Db      465 NPYAPENPKIRI08EGLHAGTGLTFTTAQDDPRYMOQNIIRYTKLSDPANMLKIDSVNGOI 524
Qy      620 FSVAPLDRREA-----GSPYRQOVVADEVGSSLSVSEFHLILMDVNDNPPRLADYGLF 675
Db      525 TTIAVLDRRESNVAANYATLFLASDNGIIPMSGTGLQITLLINDNAQOVLPQEMAFI- 583
Qy      676 FCHPLSAPGSLIFEATDDQHLFRGPH-FTFSLGSGSLQNDWEVSKINGTHARLSTRTD 734
Db      584 -CE-TPDPNSINITATLDDIDPNAGPFAFLPLSPVITIKRWMTITRLNGDPAQNLKIKF 641
Qy      735 FEEAAYVVLIRINDGNRPPLGIVSLPFTPCSC-VESGSCFRPAGHQIGIPV---GNAGV 790
Db      642 LEAGIVFVPIITITDSGNPKNSISILRVKVCQCCDSNDG-----TDVDRIVAGLGTG 694
Qy      791 ILTTLTVIGIILAVFIRIKDGKDNVESQAQSEVVP 829
Db      695 AIIAIIICITLILVLMFVVMKGRKROKQKLLIDP 733

```

RESULT 10

```

C38892
Cadenherin 4 precursor - human
N:Alternate names: R-cadherin
C:Species: Homo sapiens (man)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 20-Aug-1999
C:Accession: C38892
R/Suzuki, S.; Sano, K.; Tanihara, H.
Cell Regul. 2, 261-270, 1991
A:Title: Diversity of the cadherin family: evidence for eight new cadherins in nervous t
A:Reference number: 524305; MUID:91283540; PMID:2059658
A:Accession: C38892
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-916 <SUZ>
A:Cross-references: GB:IJ34059, NID:g506409, PIDN:AAA35627.1; PID:g506410
C:Genetics:
A:Gene: GDB:CDH4
A:Cross-references: GDB:622850
A:Map position: 16q24.1-16pter
C:Superfamily: cadherin
C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane ptc
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-166/Domain: propeptide #status predicted <PRO>
F:167-913/Product: R-cadherin #status predicted <MNT>
F:167-721/Domain: extracellular #status predicted <EXT>
F:169-274/Domain: cadherin repeat homology <CR1>
F:244-249/Region: cadherin binding #status predicted
F:277-389/Domain: cadherin repeat homology <CR2>
F:300-304/Domain: calcium binding #status predicted <CAB>
F:392-504/Domain: cadherin repeat homology <CR3>
F:507-612/Domain: cadherin repeat homology <CR4>
F:613-721/Domain: cadherin repeat homology <CR5>
F:722-753/Domain: transmembrane #status predicted <TMN>
F:754-913/Domain: intracellular #status predicted <INT>
F:870-885/Region: serine-rich
F:283,412,557,632,661,702/Binding site: carboxydrate (Asn) (covalent) #status predicted

```

```

Query Match      15.8%; Score 684.5; DB 2; Length 916;
Best Local Similarity 29.3%; Pred. No. 6e-35;
Matches 176; Conservative 106; Mismatches 288; Indels 31; Gaps 12;

Qy      239 WKAPKPEVWENSTDPHPKITQVRM---NDPGAQVSL--VDKEKLPDPFPSSID-QEGDI 292
Db      171 WVIP-PNVNPEVNSGPPQQLVTRISDKNDIPRISITGVADQPPMEVFSINSGRM 229
Qy      293 YVTPQLDRREKDAYVFAVAKDEYKPLSYPLIEHVAKDINDNPTCPSPVTVFEVOEN 352
Db      230 YVTRPMDREHSHAYHLRAHVDNMGNKVENPIDLIYIVIDMNDNHPFINQVYVNCSDRG 289
Qy      353 ERLGNSIGTLTAHDREENTANSFLNRYIEQTPKPLPMDGLFIQTYAGMLQIAKOSLKK 412
Db      290 SKRGTYVMTTANDADSTTANGWVRVYIVTQTPQSPQNMFTINSETGDIVVVAAGWDR 349

```

```

Qy      413 QTPPYNLITIEVSDKD-----FKTLCFQVQINVIDINDQIPIPEKSDYGNLTIAEDTNIG 466
Db      350 EKVOQTAVIYQATDMGNGNLTGNSNTATLITVTDVNDNSETFASIFAG--EVENSV 407
Qy      467 STLTITQATDADDEPFGSSKILYHIIKQDEBGLGVDTDPHNTGYVYIIKKPLDFTAAV 526
Db      408 TVANILVMDRDOPHSPNNAAVYRIISGPSGHFSVRTDPVTEGWTVKAVDYELNRA 467
Qy      527 SNIVFAENEPILVFGVKNVASSFAKFTLLIVDVNEAPQSCVHPQAKVSEDAVIGTKVG 586
Db      468 FMLTVVNSQAPLASGIQMSFOSTAGVYISIMDINAPFPSSNKLIRLEEGVPPGVLV 527
Qy      587 NTAQDPBEGH---DISYLRGDTFRGLKIDHVGRIFSVAPLDRREA-----GSPYRQOVA 639
Db      528 TFSVAVDPDRMQQAVYNSKLSDSPASLHINATNGQITTVAVDREGIYTGNNYEAFTLA 587
Qy      640 TEVGGSSLSVSEFHLILMDVNDNPPRLADYGLFFCHPLSAPGSLIFEATDDQHLFR 699
Db      588 ADNGIIPASGTGLQITVLLIDINDNAPELPRKAQICERPILNA---INITPAADADVHPNI 644
Qy      700 GPH-FTFSLGSGSLQNDWEVSKINGTHARLSTRTDPEERAYVVLIRINDGNRPPLGIV 758
Db      645 GPVYFELPFVPAARVARNWTTIRLNGDYAQLSLIYLEAGMVDVPIITVDSGNPPLSNTS 704
Qy      759 SLPTVFCSCVSGSCFRPAGHQIGIPVGNAGVILITLTVIGIILAV--FIRIKDKGK 816
Db      705 IIRVXKCPDDNDGCTTIG---AVALAGTGAIVAILICITLILVLMFVVMKGRK 761
Qy      817 D 817
Db      762 E 762

```

RESULT 11

```

I50116
N-cadherin precursor - zebra fish
C:Species: Brachydanio rerio (zebra fish)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jan-2000
C:Accession: I50116
R/Bitzur, S.; Kam, Z.; Geiger, B.
Dev. Dyn. 201, 121-136, 1994
A:Title: Structure and distribution of N-cadherin in developing zebrafish embryos: morp
A:Reference number: I50116; MUID:95178741; PMID:7873785
A:Accession: I50116
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-783 <BIT>
A:Cross-references: EMBL:X67648; NID:g623209; PIDN:CAA47890.1; PID:g623210
C:Superfamily: cadherin
F:146-258/Domain: cadherin repeat homology <CDH>

```

```

Query Match      15.7%; Score 677; DB 2; Length 783;
Best Local Similarity 30.1%; Pred. No. 1.4e-34;
Matches 187; Conservative 98; Mismatches 288; Indels 48; Gaps 15;

Qy      239 WKAPKPEVWENSTDPHP---IKITQVRKMDGCAQVSLV--DKEKLPDPFPSSID-QEGDI 292
Db      37 WVIP-PNVNPEVNSRQKQFPELVKIQSDKXSLTRYSVTPGADQNPFGITFIIDPISGL 95
Qy      293 YVTPQLDRREKDAYVFAVAKDEYKPLSYPLIEHVAKDINDNPTCPSPVTVFEVOEN 352
Db      96 SVTKPLDRHIIHPRFLRAHVDINGNQMNPIIDIIINVIDMNDNHPFTHQINGVYDGG 155
Qy      353 ERLGNSIGTLTAHDREENTANSFLNRYIEQTPKPLPMDGLFIQTYAGMLQIAKOSLKK 412
Db      156 AKRGTFVMTVTSQDKDPDPTANGMLRYKILSTQTPSPSSNMFTINNKTKIITVAAGLDR 215
Qy      413 QTPPYNLITIEVSDKD-----FKTLCFQVQINVIDINDQIPIPEKSDYGNLTIAEDTNIG 466
Db      216 EKVPQYTLITIQATDMGNGNLTGNSNTATVAILDVNDNAPFTRETFEGVEVNNRVNY 275
Qy      467 STLTITQATDADDEPFGSSKILYHIIKQDEBGLGVDTDPHNTGYVYIIKKPLDFTAAV 526

```

```

Db      276 VTULTV--TDKDEGTAPMAAVYRIISGDPGRSIPITDPTVNEGLVTVVKKPVDVFNRRS 333
      527 SNIVFAKNEPEPLVFGYKKNASSFAKTLTYTVNEAPPSQHFQAKVSEDAVIGIKVG 586
      334 FMLTVADVADNEVPLASGHTRTQSTATVTSIRVIDNEBPNPDPNKQIKLEEGFPQWMLT 393
      587 NVTA-KQPEGL---DISYSLRGDTRGMLKIDHVTGELFSVAPLDREA---GSPYRQVV 638
      394 TPTAHDPDRYMOQTISYSKLYDPANLEIDPNNGRSTIAVLDRSPYKANNLYNTFM 453
      639 ATTEVGSLSVSEFHLILMDVNDNPPRLAKDYTLGFCHPLSA-----PGSLIFRAT 691
      454 ASDNGVPRASGTGLQIYLLDINDNARV-----PQGEAVGCEBPANINITAV 503
      692 DDDHGLFRGPHFTSLSS--GSLONDEVSKINGTHARLS-TRHTDFEERAVVLIRINDG 749
      504 DGDINPNVAGP-YAFELNRPBIDIRNWTLTRISGDHQLSLKISYLSGIEYELFISTDS 562
      750 GRPLEGIVSLPTVFCSC-VEGSCFRPAGHOTGIPTVGMAYGILLTLVLVIGITLAVFI 808
      563 GNLFMSNTTYLRIRVCCQDHHRGCV---DMERIMAGLGITGAIILICIIILVLVM 618
      809 RIKKDKGKNVESAQASEVKP 829
      619 FVMMKKRRDKERQAKOLLIDP 639

```

RESULT 12

```

B55363
desmocolin, type 4, short form precursor - human
C/Species: Homo sapiens (man)
C/Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 23-May-1997
C/Accession: B55363
R/Kawamura, K.; Watanabe, K.; Suzuki, T.; Yamakawa, T.; Kamiyama, T.; Nakagawa, H.; Tsui
J. Biol. Chem. 269, 26295-26302, 1994
A/Title: cDNA cloning and expression of a novel human desmocolin.
A/Reference number: A55363; MUID:95014464; PMID:7929347
A/Accession: B55363
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-826 <KAW>
A/Cross-references: GB:D17427
C/Superfamily: cadherin; cadherin repeat homology
C/Keywords: alternative splicing; transmembrane protein
F/138-243/Domain: cadherin repeat homology <CR1>
F/246-355/Domain: cadherin repeat homology <CR2>

```

```

Query Match      15.5%; Score 671.5; DB 2; Length 826;
Best Local Similarity 30.8%; Pred. No. 3.3e-34;
Matches 187; Conservative 106; Mismatches 263; Indels 51; Gaps 22;

      239 WKAPKEVEMVNSTDPPIKITQVRWMDPGAQYSL-----VDKEKLPRFPISIDQEG 290
      137 W-APICSMQENSIGPPLFLQOVE-SDAQNVTVFYISGRGVDEKPELNFYIERD-TG 193
      291 DIYVTOPLDREKDAY--VFYAVAKDEYKGLSYPLEIHVKVQDINDNPTCSPVYVE 348
      194 NLCTRPVDEEYVDVLAIVASTADGYSADL--PLPLPRVDEEDNHNHVFTEALYNFE 251
      349 VOENBERGNSIGTLTAHDBREENTANSFLNRYIVEQPKLPMDGLFLIQTAGMLQAKQ 408
      252 VLESSRRTYGVGVCATDRDEPDMTRKYSILQOTPRSP--GLFSVHSTGVITTVSH 309
      409 SLKKDTPQYNLTIEVSDKPK-----TLCFVQVINVDINDQIPFEKSDYGNLTIAE 461
      310 YLDREVVDKYSILMKVQDMQGFGLIGTSTCI--ITVDSNDNAPFRQNAVE--AFVE 365
      462 DTNIGSTILTIOATDADEPFTGSSKILYHIKDSBGLGVDTPDHTNTGVIILKPLDF 521
      366 ENAFNVEILRIPIEDKDLINTANMRVNFILIKGENGHFKISTDKETNEGVLSVVKPLNY 425
      522 ETAAVSNIVFAKNEPEPLVFGV-KYNASSFAKTLTYTVNEAPPSQHFQAKVSEDA 580

```

```

      426 EENRQVNLIEGVNNEAPFARDIPRYALNRALVTVHVRDLDEDEPECTPAQOYVRIKENLA 485
      581 IGTQVGNVTAQDE---GLDISYSLRGDTRGMLKIDHVTGELFSVAPLDREAQSP-----Y 633
      486 VSGIKNGYKAYDENENNGLRKJLHDPGQWTTIDEISGSIITSKILBREVTPOKELY 545
      634 RVGVNATEVGSLSVSEFHLILMDVNDNPPRLAKDYTLGFCHPLSAAGSLIFRATD 693
      546 NITVALAIDDDRECTGTIAVN--IEDVNDNPELIDQY--VVICRKMGYTDIL--AVDP 599
      694 DQHLFRGPHFTSLSSGS--LONDEVSKINGTHARLS-TRHTDFEERAVVLIRINDG 750
      600 DEBVHGAAP-FYFSLPNTSPEISRLMSLTLYVNDPAARLSQKNAQFOE--YTIPIYKD-- 654
      751 RPLLEGIVSLPTVFCSCVSGSCFRPAGHOTGIPTVGMAY-GILLTLVLVIGITLAVFI 809
      655 RAGQAAKLLRVNLCCEHTPQCRATRSSTIGVLKGMALAILLILFSLVLTIVCGV 714
      810 IKKDKGK 816
      715 FGATKCK 721

```

RESULT 13

```

A55363
desmocolin, type 4, long form precursor - human
C/Species: Homo sapiens (man)
C/Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 20-Jun-2000
C/Accession: A55363
R/Kawamura, K.; Watanabe, K.; Suzuki, T.; Yamakawa, T.; Kamiyama, T.; Nakagawa, H.; Tsui
J. Biol. Chem. 269, 26295-26302, 1994
A/Title: cDNA cloning and expression of a novel human desmocolin.
A/Reference number: A55363; MUID:95014464; PMID:7929347
A/Accession: A55363
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-896 <KAW>
A/Cross-references: GB:D17427; NID:9639672; PIDN:BAA04249.1; PID:9685219
C/Superfamily: cadherin; cadherin repeat homology
C/Keywords: alternative splicing; transmembrane protein
F/138-243/Domain: cadherin repeat homology <CR1>
F/246-355/Domain: cadherin repeat homology <CR2>

```

```

Query Match      15.5%; Score 671.5; DB 2; Length 896;
Best Local Similarity 30.8%; Pred. No. 3.8e-34;
Matches 187; Conservative 106; Mismatches 263; Indels 51; Gaps 22;

      239 WKAPKEVEMVNSTDPPIKITQVRWMDPGAQYSL-----VDKEKLPRFPISIDQEG 290
      137 W-APICSMQENSIGPPLFLQOVE-SDAQNVTVFYISGRGVDEKPELNFYIERD-TG 193
      291 DIYVTOPLDREKDAY--VFYAVAKDEYKGLSYPLEIHVKVQDINDNPTCSPVYVE 348
      194 NLCTRPVDEEYVDVLAIVASTADGYSADL--PLPLPRVDEEDNHNHVFTEALYNFE 251
      349 VOENBERGNSIGTLTAHDBREENTANSFLNRYIVEQPKLPMDGLFLIQTAGMLQAKQ 408
      252 VLESSRRTYGVGVCATDRDEPDMTRKYSILQOTPRSP--GLFSVHSTGVITTVSH 309
      409 SLKKDTPQYNLTIEVSDKPK-----TLCFVQVINVDINDQIPFEKSDYGNLTIAE 461
      310 YLDREVVDKYSILMKVQDMQGFGLIGTSTCI--ITVDSNDNAPFRQNAVE--AFVE 365
      462 DTNIGSTILTIOATDADEPFTGSSKILYHIKDSBGLGVDTPDHTNTGVIILKPLDF 521
      366 ENAFNVEILRIPIEDKDLINTANMRVNFILIKGENGHFKISTDKETNEGVLSVVKPLNY 425
      522 ETAAVSNIVFAKNEPEPLVFGV-KYNASSFAKTLTYTVNEAPPSQHFQAKVSEDA 580
      426 EENRQVNLIEGVNNEAPFARDIPRYALNRALVTVHVRDLDEDEPECTPAQOYVRIKENLA 485
      581 IGTQVGNVTAQDE---GLDISYSLRGDTRGMLKIDHVTGELFSVAPLDREAQSP-----Y 633

```

```

Db      486 VGSKINGYKAYDENRNGNGLRYKCLADPKGWLITIDEISGSIITSKILIDREVEBPKNELY 545
      634 RGVVNAVEVSGSSLSVSEFHLIMDVNDNPPRLAKOYGLFCHPLSAGSLIFETDD 633
      546 NITVLALDKDRCSTGLTAVN--IEDVNDNPPRLQGY--VVICPKPMGYTDL--AVDP 599
      694 DQHLFRGPHFTFSGSGS--LQNDWEVSKINGTHARLS-TRHDPFEERAYVVLIRINDG 750
      600 DEPVHGP-PYFSLPNTSPSISRLMSLTKNNDTRALSTYKMGFGP--YTIPIYVD-- 654
      751 RPLEGIVSLPYTFSCVSGSCFRPAGHGQGIPTVNAV-GILITLLVIGIILAVFIR 809
      655 RAGQAMTKLRVNLCEHTHPTQCRATSRSTGVILGKMAILAILLIGALLFSLVLTAVGV 714
      810 IKKDKGK 816
      715 FGATKCK 721
      Db

RESULT 14
LHUBD
desmocollin 3b precursor - human
N/Alternate names: desmosomal glycoprotein III
C/Species: Homo sapiens (man)
C/Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 07-May-1999
C/Accession: A40390; S16464
R/Parker, A.E.; Wheeler, G.N.; Arneemann, J.; Pidsley, S.C.; Ataliotis, P.; Thomas, C.L.
J. Biol. Chem. 266, 10438-10445, 1991
A/Title: Desmosomal glycoproteins II and III. Cadherin-like junctional molecules generated
A/Reference number: A40390; MUID:91244819; PMID:2037591
A/Accession: A40390
A/Molecule type: mRNA
A/Residues: 1-847 <PAR>
A/Cross-references: GB:X56807
A/Note: It is uncertain whether Met-1 is the initiator or whether translation is initiated
R/Buxton, R.S.; Cowin, P.; Franke, W.W.; Garrod, D.R.; Green, K.J.; King, I.A.; Koch, P.
J. Cell Biol. 121, 481-483, 1993
A/Title: Nomenclature of the desmosomal cadherins.
A/Reference number: A43032; MUID:93252984; PMID:8486729
A/Contents: annotation; nomenclature
C/Genetics:
A/Gene: GDB:DSC3; DSC2; DSC1; DS
A/Cross-references: GDB:126552; OMIM:600271
A/Map position: 18q12.1-18q12.1
C/Keywords: cadherin; cadherin repeat homology
F/1-28/Domain: signal sequence #status predicted <PRO>
F/29-135/Domain: propeptide #status predicted <PRO>
F/136-847/Product: desmocollin 3b #status predicted <EXT>
F/136-695/Domain: extracellular #status predicted <EXT>
F/138-243/Domain: cadherin repeat homology <CR1>
F/246-355/Domain: cadherin repeat homology <CR2>
F/358-471/Domain: cadherin repeat homology <CR3>
F/474-577/Domain: cadherin repeat homology <CR4>
F/578-680/Domain: cadherin repeat homology <CR5>
F/696-718/Domain: transmembrane #status predicted <TM>
F/719-847/Domain: intracellular #status predicted <INT>
F/166,392,546,629/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      15.4%; Score 667.5; DB 1; Length 847;
Best local similarity 30.9%; Pred. No. 6,2e-34;
Matches 186; Conservative 104; Mismatches 253; Indels 59; Gaps 21;

      QY      239 WKAPKVEVENVSTDPHPIKITOVRMNDPAQYSL-----VDKEKLPFPSPIDOE- 289
      DB      137 W-APIPSMLNLSIGPPPLFLQOYQ-SDTAQNTIYYSINGPVDOE--FRNLFYVERDT 192
      QY      290 GDLYVQPLDREEDAVFAVAKDEYKSLVPLELHVAVKQINDNPPCPSPVYFEV 349
      DB      193 GNIYCRPVRQEGYSEETIAFAATTPGYPPELPLIITIEDENDYAFITSETTYFTI 252
      QY      350 QENRNLGNSIGTLTAHDREENTANSFLNRYIVEQTPKLPMDGLFIOTYAGMLQIAKOS 409

```

```

      253 FENCRVGTIVGQVCANDKDEPDMTRLAKSLIGQP--PSPTLFAMHPTTGVIITTSQ 310
      410 LKKQDPQVNLITVSDKD-----FKTLGFQVNVNDINDQPIFEKSYGNLTAEEDN 464
      311 LDRELIDKYOQKIKVQDMQGYFGLOTSTCIINDVDNDHLPFTRTS--VTSVEENT 368
      465 IGSITLITQATDADEPFGSSKILYHIKQDSGRGLGVTDPPHTNGVYLIKPLDPE-- 522
      369 VDVEILRYVEDDLVNTANRANNTIILKNGENGNKIYTDKATNGSVLCVQKPLVEEK 428
      523 -----TAAVSNIVFAE-NPEELVGVKYNASSFAKFTLLVTVDNAEPQSFQHVQAKY 575
      429 QQWILQIGVNNRPAFRESAPR-----SAMSTATVTVNVEDQEGECNPIQTVNM 480
      576 SEDVALGTQVANTADPE---GLDYSYLRGTRWKLIDHYTGSIIFSAPLDREA--- 629
      481 KENAEVGTTSNGYKADPPTRSSSGIRYKCLTDPTMWTIDNTGSIKVFRLDRBAETI 540
      630 -GSPRYOVVATEVSGSSLSVSEFHLIMDVNDNPPRLAKOYGLFCHPLSAPGLP 688
      541 KNGIYVITLASDQGRCTGCT--LGITLQDVNDNSPPIPK--KVIYICKPTMSSAIV- 595
      689 EATDDQHLFRGPHFTFSGSGS--LQNDWEVSKINGTHARLSTRHDPFEERAYVVLIRI 746
      596 -AVDPDEPI-HGPPDFSLSESTSEYQRMWRMLKAINDTARLSYQ--DPPFGSYVPITV 652
      747 NDGRRPLEGIVSLPYTFSCV-EGSCFRPAGHGQGIPTVNAVAGILITLLVIGIILAV 805
      653 RD--RLGMSVSVSLDVTLCDTTENDCTRVDPRIGGGVQLGKMAILAILLIGALLPFI 710
      806 VF 807
      711 LP 712
      Db

RESULT 15
LHUBD
desmocollin 3a precursor - human
N/Alternate names: desmosomal glycoprotein II
C/Species: Homo sapiens (man)
C/Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 07-May-1999
C/Accession: B40390; S16465
R/Parker, A.E.; Wheeler, G.N.; Arneemann, J.; Pidsley, S.C.; Ataliotis, P.; Thomas, C.L.
J. Biol. Chem. 266, 10438-10445, 1991
A/Title: Desmosomal glycoproteins II and III. Cadherin-like junctional molecules genera
A/Reference number: A40390; MUID:91244819; PMID:2037591
A/Accession: B40390
A/Molecule type: mRNA
A/Residues: 1-901 <PAR>
A/Cross-references: GB:X56807
A/Note: It is uncertain whether Met-1 is the initiator or whether translation is initia
R/Buxton, R.S.; Cowin, P.; Franke, W.W.; Garrod, D.R.; Green, K.J.; King, I.A.; Koch, P.
J. Cell Biol. 121, 481-483, 1993
A/Title: Nomenclature of the desmosomal cadherins.
A/Reference number: A43032; MUID:93252984; PMID:8486729
A/Contents: annotation; nomenclature
C/Genetics:
A/Gene: GDB:DSC3; DSC2; DSC1; DS
A/Cross-references: GDB:126552; OMIM:600271
A/Map position: 18q12.1-18q12.1
C/Keywords: cadherin; cadherin repeat homology
F/1-28/Domain: signal sequence #status predicted <PRO>
F/29-135/Domain: propeptide #status predicted <PRO>
F/136-901/Product: desmocollin 3a #status predicted <EXT>
F/138-243/Domain: cadherin repeat homology <CR1>
F/246-355/Domain: cadherin repeat homology <CR2>
F/358-471/Domain: cadherin repeat homology <CR3>
F/474-577/Domain: cadherin repeat homology <CR4>
F/578-680/Domain: cadherin repeat homology <CR5>
F/696-718/Domain: transmembrane #status predicted <TM>
F/696-718/Domain: transmembrane #status predicted <TM>

```

F:719-901/Domain: intracellular #status predicted <INT>
F:166 392,546,629/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:664/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match

15.4%; Score 667.5; DB 1; Length 901;

Best Local Similarity 30.9%; Pred. No. 6,9e-34;
Matches 186; Conservative 104; Mismatches 253; Indels 59; Gaps 21;

```
QY 239 WKAKPEVMENSTDPHPKIKITQVRMNDPGAQYSL-----VDKEKLPFPSPIDOE- 289
DB 137 W-ADIPCSMTLNSLGPPLFLQOVQ-SDTAQNITYISIRGPVQDE--PRLNLFYERDT 192
QY 290 GDIIVTQPLDREKDAVYFAVAKDEYKPLSYPLBIHVKKVDINDNPPTCSPVTVFEV 349
DB 193 GNLVCTRPVDRQYSEFELIAFATTPDGYTPELPPLIKIEDENDNVPITFEETYTFTI 252
QY 350 QENRRLGNSIGTLTAHDBRENTANSFLNRIVEQTPKLPMDGLFLIQTATGMIQLAKOS 409
DB 253 FENCRVGTWVQVQCATKDEPDTHMLKYSIIIGQVP--PSPTLFSMHPTTGVITTSQ 310
QY 410 LKKQDTQVYNLTIEVSKD-----FKTLCFQVINVIDINDQIPIFEKSDVGNLTLEADTN 464
DB 311 LDRRLIKYQKIKIVQMDQYFGLQTTSTCIINDVDNHLPTFTRTSY--VTSVEBNT 368
QY 465 IGSTILITATDADEPFTGSSKILYHIKDSERLGVTDPTNTGYVILKKPLDPE-- 522
DB 369 VDVRIILKVTVEDKDLVNTAMNRYTILKGNENGFKYTDKATNEGVLGVKRLNVEEK 428
QY 523 -----TAVSNIVFKAE-NPEPLVFGVKYNASSPAKFTLIIVDVNEAPQSFQHVFOAKV 575
DB 429 QQMILQIGVNVAEAFSEASPR-----SAMSTATVTVNVEDQDEGPECNPEIQTVRM 480
QY 576 SEDVAIGTKVGNVTAKEPE---GLDISLNGDTRGMKIDHVTGEIIFSVAPLDREA--- 629
DB 481 KENAEVGTTSNGYAYADETSSSGIRYKLLDPTGWTTIDENTGSIKVFERSLDREAFTI 540
QY 630 -GSPYRVQVVADEVGSSLSVSEFHLILMDVNDNPPRLADYTGLEFCHPLSAPGSLIF 688
DB 541 KNGIYNITVLASDGGRTCTGT--LGIILQDVNDNSPPIPK--KTVIICKPTMSAEIV- 595
QY 689 EATDDDHLEFRGPHFTSLGSGS--LQNDMEVSKINGTHARLSTRHTDFFERAYVVLIRI 746
DB 596 -AVDPDEPI-HGPPDFEFLSSTSEVQRMRLKAINDTAARLSYQN--DPPFGSYVVPITV 652
QY 747 NDGGRPLEGIVSLPVTFCSGV-EGSCFRPAGHOTGIPYGMANGILLTLLVGIILAV 805
DB 653 RD--RLQMSVTSLSLVTLCDCTITENDCTHRVDPRIQGGGVQLGKMAITAILLALFCTI 710
QY 806 VF 807
DB 711 LF 712
```

Search completed: June 22, 2003, 01:08:45
Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2003, 01:04:45 ; Search time 62 Seconds
(without alignments)
1452.066 Million cell updates/sec

Title: US-10-025-380-1081

Perfect score: 4321
Sequence: 1 MIQAHLSICLIMLYATG.....DKGKNVSAQAEVPLRS 832

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/2/pubppaa/PTCT_NEW_PUB pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB pep:*
6: /cgn2_6/ptodata/2/pubppaa/PTCT_PUBCOMB pep:*
7: /cgn2_6/ptodata/2/pubppaa/PTCT_PUBCOMB pep:*
8: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/2/pubppaa/US05_PUBCOMB pep:*
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4321	100.0	832	9	US-10-025-380-1081
2	4321	100.0	832	10	US-09-922-217-1081
3	4321	100.0	832	10	US-09-833-263-1081
4	4208	97.4	958	9	US-10-025-380-1087
5	4208	97.4	958	10	US-09-922-217-1087
6	4208	97.4	958	10	US-09-833-263-1087
7	908.5	21.0	807	9	US-10-063-547-98
8	908.5	21.0	807	9	US-10-063-516-98
9	908.5	21.0	807	9	US-10-063-502-98
10	908.5	21.0	807	9	US-10-227-884-188
11	908.5	21.0	807	9	US-10-230-163-188
12	908.5	21.0	807	9	US-10-006-856A-229
13	908.5	21.0	807	9	US-10-218-631-188
14	908.5	21.0	807	9	US-10-230-338-188
15	908.5	21.0	807	9	US-10-063-518-98
16	908.5	21.0	807	9	US-10-063-598-98
17	908.5	21.0	807	9	US-10-227-693-98
18	908.5	21.0	807	9	US-10-230-414-188
19	908.5	21.0	807	9	US-10-006-818A-229

20	908.5	21.0	807	9	US-10-015-393A-229	Sequence 229, App
21	908.5	21.0	807	9	US-10-063-567-98	Sequence 98, App1
22	908.5	21.0	807	9	US-10-216-159A-188	Sequence 188, App
23	908.5	21.0	807	9	US-09-846-374-229	Sequence 229, App
24	908.5	21.0	807	9	US-10-012-121A-229	Sequence 229, App
25	908.5	21.0	807	9	US-10-015-869A-229	Sequence 229, App
26	908.5	21.0	807	9	US-10-063-538-98	Sequence 98, App1
27	908.5	21.0	807	9	US-10-218-849-188	Sequence 188, App
28	908.5	21.0	807	9	US-10-227-873-188	Sequence 188, App
29	908.5	21.0	807	9	US-10-227-883-188	Sequence 188, App
30	908.5	21.0	807	9	US-10-063-599-98	Sequence 98, App1
31	908.5	21.0	807	9	US-10-219-076-188	Sequence 188, App
32	908.5	21.0	807	9	US-10-230-434-188	Sequence 188, App
33	908.5	21.0	807	9	US-10-006-116A-229	Sequence 229, App
34	908.5	21.0	807	9	US-10-006-117A-229	Sequence 229, App
35	908.5	21.0	807	9	US-10-013-913A-229	Sequence 229, App
36	908.5	21.0	807	9	US-10-017-527A-229	Sequence 229, App
37	908.5	21.0	807	9	US-10-063-595-98	Sequence 98, App1
38	908.5	21.0	807	9	US-10-219-003-188	Sequence 188, App
39	908.5	21.0	807	9	US-10-219-075-188	Sequence 188, App
40	908.5	21.0	807	9	US-10-219-464-188	Sequence 188, App
41	908.5	21.0	807	9	US-10-219-466-188	Sequence 188, App
42	908.5	21.0	807	9	US-10-219-479-188	Sequence 188, App
43	908.5	21.0	807	9	US-10-219-481-188	Sequence 188, App
44	908.5	21.0	807	9	US-10-230-260-188	Sequence 188, App
45	908.5	21.0	807	9	US-10-232-231-188	Sequence 188, App

ALIGNMENTS

RESULT 1
US-10-025-380-1081
Sequence 1081, Application US/10025380
Publication No. US20020182191A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Stolk, John A.
APPLICANT: Wang, Tongtong
APPLICANT: Jiang, Yugu
APPLICANT: Smith, Carole L.
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Skeiky, Yasir A. W.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick Thomas S.
APPLICANT: Carter, Darrick
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
FILE REFERENCE: 210121.471C14
CURRENT APPLICATION NUMBER: US/10/025,380
NUMBER OF SEQ ID NOS: 1129
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1081
LENGTH: 832
TYPE: PRT
ORGANISM: Homo sapiens
US-10-025-380-1081

Query Match 100.0%; Score 4321; DB 9; Length 832;
Best Local Similarity 100.0%; Pred. No. 1.6e-315;

Matches 832; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIQAHLSICLIMLYATGYGEGKFGSLKPMFTSIYEGEPSSQITTFQKPNPAVTF 60
DB 1 MIQAHLSICLIMLYATGYGEGKFGSLKPMFTSIYEGEPSSQITTFQKPNPAVTF 60

```

QY 61 ELTGETDNIFVIEREGLLYYNRALDRETRSTHNLQVAALDANGIIVEGPVITIEVDIN 120
DB 61 ELTGETDNIFVIEREGLLYYNRALDRETRSTHNLQVAALDANGIIVEGPVITIEVDIN 120
QY 121 DNRPTFLOSKYEGSVRQNSRPGKPELYVNATLDDPATNGOLYYQIVITQPMINMYF 180
DB 121 DNRPTFLOSKYEGSVRQNSRPGKPELYVNATLDDPATNGOLYYQIVITQPMINMYF 180
QY 181 QINNKTAISLTREGSOELNPAKNPSYNLVSVKMGQSGNSFSDTTSVDIIVTENIMK 240
DB 181 QINNKTAISLTREGSOELNPAKNPSYNLVSVKMGQSGNSFSDTTSVDIIVTENIMK 240
QY 241 APRKVENWENSTDHPKIKITQVRMNDGAGYSLVDEKELPRPFPSIDQEGDIYVTOPLDR 300
DB 241 APRKVENWENSTDHPKIKITQVRMNDGAGYSLVDEKELPRPFPSIDQEGDIYVTOPLDR 300
QY 301 EEXDAYVFAVAKDEYKPLSYPLEIHVKVDINDNPTCPSPVTFVEVOENERLGNISIG 360
DB 301 EEXDAYVFAVAKDEYKPLSYPLEIHVKVDINDNPTCPSPVTFVEVOENERLGNISIG 360
QY 361 TLTAHDEDEENTANSFLNRYIVEQTPKLPMDGFLQTYAGMLQAKOSLKKODTPOYNL 420
DB 361 TLTAHDEDEENTANSFLNRYIVEQTPKLPMDGFLQTYAGMLQAKOSLKKODTPOYNL 420
QY 421 TIEVSDKDFKLCFVOQINVDINDQIPIFEKSDYGNLTAEADNIGSTILTITQATDADEP 480
DB 421 TIEVSDKDFKLCFVOQINVDINDQIPIFEKSDYGNLTAEADNIGSTILTITQATDADEP 480
QY 481 FTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVIKKPLDEFTAAVSNIVFKAENPEPLV 540
DB 481 FTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVIKKPLDEFTAAVSNIVFKAENPEPLV 540
QY 541 FGKYNASSPAKFTLLVTDVNEAPOFSQHVFOAKVSEDAVIGTKVGNVTADEPGLDISY 600
DB 541 FGKYNASSPAKFTLLVTDVNEAPOFSQHVFOAKVSEDAVIGTKVGNVTADEPGLDISY 600
QY 601 SLRGDTRGMLKIDHVTGEIFSVAPLDRAGSPRYOVVATEVGGSSLSVSEFHLILMDV 660
DB 601 SLRGDTRGMLKIDHVTGEIFSVAPLDRAGSPRYOVVATEVGGSSLSVSEFHLILMDV 660
QY 661 NDNPRLAKDYTGFLFCHPLSAPGSLIFEATDDOHLFRGPHFTFSLGSGSLQNDWEVSK 720
DB 661 NDNPRLAKDYTGFLFCHPLSAPGSLIFEATDDOHLFRGPHFTFSLGSGSLQNDWEVSK 720
QY 721 INGTARLSTRHTDPEERAYVVLIRINDGRPLREGIVSLPVTFCSCEGSCFRPAGHOT 780
DB 721 INGTARLSTRHTDPEERAYVVLIRINDGRPLREGIVSLPVTFCSCEGSCFRPAGHOT 780
QY 781 GIPVGAAGVILTTLLVIGIILAVVFIRIKKDKGKNVESAOASEVKKPLRS 832
DB 781 GIPVGAAGVILTTLLVIGIILAVVFIRIKKDKGKNVESAOASEVKKPLRS 832

```

RESULT 2

```

US-09-922-217-1081
; Sequence 1081, Application US/09922217
; Patent No. US2002007641A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secretat, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stoik, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13

```

```

; CURRENT APPLICATION NUMBER: US/09/922,217
; NUMBER FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1081
; LENGTH: 832
; TYPE: PR
; ORGANISM: Homo sapiens
US-09-922-217-1081

```

```

Query Match 100.0%; Score 4321; DB 10; Length 832;
Best Local Similarity 100.0%; Pred. No. 1,6e-315;
Matches 832; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MIQAHLSICLMLYLANGYGOEGKSPGKMTESIYEGQSPQIIFQFKANPPAYF 60
DB 1 MIQAHLSICLMLYLANGYGOEGKSPGKMTESIYEGQSPQIIFQFKANPPAYF 60
QY 61 ELTGETDNIFVIEREGLLYYNRALDRETRSTHNLQVAALDANGIIVEGPVITIEVDIN 120
DB 61 ELTGETDNIFVIEREGLLYYNRALDRETRSTHNLQVAALDANGIIVEGPVITIEVDIN 120
QY 121 DNRPTFLOSKYEGSVRQNSRPGKPELYVNATLDDPATNGOLYYQIVITQPMINMYF 180
DB 121 DNRPTFLOSKYEGSVRQNSRPGKPELYVNATLDDPATNGOLYYQIVITQPMINMYF 180
QY 181 QINNKTAISLTREGSOELNPAKNPSYNLVSVKMGQSGNSFSDTTSVDIIVTENIMK 240
DB 181 QINNKTAISLTREGSOELNPAKNPSYNLVSVKMGQSGNSFSDTTSVDIIVTENIMK 240
QY 241 APRKVENWENSTDHPKIKITQVRMNDGAGYSLVDEKELPRPFPSIDQEGDIYVTOPLDR 300
DB 241 APRKVENWENSTDHPKIKITQVRMNDGAGYSLVDEKELPRPFPSIDQEGDIYVTOPLDR 300
QY 301 EEXDAYVFAVAKDEYKPLSYPLEIHVKVDINDNPTCPSPVTFVEVOENERLGNISIG 360
DB 301 EEXDAYVFAVAKDEYKPLSYPLEIHVKVDINDNPTCPSPVTFVEVOENERLGNISIG 360
QY 361 TLTAHDEDEENTANSFLNRYIVEQTPKLPMDGFLQTYAGMLQAKOSLKKODTPOYNL 420
DB 361 TLTAHDEDEENTANSFLNRYIVEQTPKLPMDGFLQTYAGMLQAKOSLKKODTPOYNL 420
QY 421 TIEVSDKDFKLCFVOQINVDINDQIPIFEKSDYGNLTAEADNIGSTILTITQATDADEP 480
DB 421 TIEVSDKDFKLCFVOQINVDINDQIPIFEKSDYGNLTAEADNIGSTILTITQATDADEP 480
QY 481 FTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVIKKPLDEFTAAVSNIVFKAENPEPLV 540
DB 481 FTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVIKKPLDEFTAAVSNIVFKAENPEPLV 540
QY 541 FGKYNASSPAKFTLLVTDVNEAPOFSQHVFOAKVSEDAVIGTKVGNVTADEPGLDISY 600
DB 541 FGKYNASSPAKFTLLVTDVNEAPOFSQHVFOAKVSEDAVIGTKVGNVTADEPGLDISY 600
QY 601 SLRGDTRGMLKIDHVTGEIFSVAPLDRAGSPRYOVVATEVGGSSLSVSEFHLILMDV 660
DB 601 SLRGDTRGMLKIDHVTGEIFSVAPLDRAGSPRYOVVATEVGGSSLSVSEFHLILMDV 660
QY 661 NDNPRLAKDYTGFLFCHPLSAPGSLIFEATDDOHLFRGPHFTFSLGSGSLQNDWEVSK 720
DB 661 NDNPRLAKDYTGFLFCHPLSAPGSLIFEATDDOHLFRGPHFTFSLGSGSLQNDWEVSK 720
QY 721 INGTARLSTRHTDPEERAYVVLIRINDGRPLREGIVSLPVTFCSCEGSCFRPAGHOT 780
DB 721 INGTARLSTRHTDPEERAYVVLIRINDGRPLREGIVSLPVTFCSCEGSCFRPAGHOT 780
QY 781 GIPVGAAGVILTTLLVIGIILAVVFIRIKKDKGKNVESAOASEVKKPLRS 832
DB 781 GIPVGAAGVILTTLLVIGIILAVVFIRIKKDKGKNVESAOASEVKKPLRS 832

```

RESULT 3

```

US-09-933-263-1081

```


Sequence 1081, Application US/09833263
Patent No. US20020110547A1
GENERAL INFORMATION:
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Stolk, John A.
APPLICANT: Meagher, Madeleine J.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
FILE REFERENCE: 210121.471C12
CURRENT APPLICATION NUMBER: US/09/833,263
NUMBER OF SEQ ID NOS: 1093
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1081
LENGTH: 832
TYPE: PR1
ORGANISM: Homo sapiens
US-09-833-263-1081

Query Match 100.0%; Score 4321; DB 10; Length 832;
Best Local Similarity 100.0%; Pred. No. 1,6e-315;
Matches 832; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIIQAHLSLCLMLYATGYGSGKSGPLKMTFSIYEGQEPSQIIIFQKPNPAVTF 60
DB 1 MIIQAHLSLCLMLYATGYGSGKSGPLKMTFSIYEGQEPSQIIIFQKPNPAVTF 60
QY 61 ELTGEDNITFIREGGLLYNRALDRETRSTHNOVALDANGIYEGPPIITEVDIN 120
DB 61 ELTGEDNITFIREGGLLYNRALDRETRSTHNOVALDANGIYEGPPIITEVDIN 120
QY 121 DNRPTFLOSKEYGSRONSRRKPLVYNATDLDPPATNGOLYYQIVOLPMINMYF 180
DB 121 DNRPTFLOSKEYGSRONSRRKPLVYNATDLDPPATNGOLYYQIVOLPMINMYF 180
QY 181 QINNKTAISLTREGSELNPAKNPSYNLVISVDMGQSGENSESDTTSVDIIVTENIMK 240
DB 181 QINNKTAISLTREGSELNPAKNPSYNLVISVDMGQSGENSESDTTSVDIIVTENIMK 240
QY 241 APPRVENENSTDPHPKIKITQVRWNDGAQYSLVDKELRPFPSIDQEBDIYVTOPLDR 300
DB 241 APPRVENENSTDPHPKIKITQVRWNDGAQYSLVDKELRPFPSIDQEBDIYVTOPLDR 300
QY 301 EEDKAYFVAVADEYKPLSYPLEIHVKYKNDINDNPTCSPVTVFEVQENERLGNISIG 360
DB 301 EEDKAYFVAVADEYKPLSYPLEIHVKYKNDINDNPTCSPVTVFEVQENERLGNISIG 360
QY 361 TLTFAHDEBENTANSFLNRYIVEQTPKLPMDGFLIQTIVAGMLQAKOSLKKODTPQYNL 420
DB 361 TLTFAHDEBENTANSFLNRYIVEQTPKLPMDGFLIQTIVAGMLQAKOSLKKODTPQYNL 420
QY 421 TIEVSDQFTKLCFVOQINVIDINDQIPFEKSDYGNITLADBNIGSTIITLQATDDEP 480
DB 421 TIEVSDQFTKLCFVOQINVIDINDQIPFEKSDYGNITLADBNIGSTIITLQATDDEP 480
QY 481 FTGSSKLIYHIIKGDSEGRIGVDDPHTNTGYVITKPLDEPTAASNIYFKAKENPRV 540
DB 481 FTGSSKLIYHIIKGDSEGRIGVDDPHTNTGYVITKPLDEPTAASNIYFKAKENPRV 540
QY 541 FGKYNASFAKFTLITVDVNEAPQFSOHVFOAKVSEDAVIGTKVGNVTADEGLDISY 600
DB 541 FGKYNASFAKFTLITVDVNEAPQFSOHVFOAKVSEDAVIGTKVGNVTADEGLDISY 600
QY 601 SLGSDTRGMLKIDHVTGEISVAPLDBEAGSPYRVQVATEVGGSSISVSEPHLIMDY 660
DB 601 SLGSDTRGMLKIDHVTGEISVAPLDBEAGSPYRVQVATEVGGSSISVSEPHLIMDY 660
QY 661 NDNPRLAKDYTGGLFCFHPISAPGSLIFBATDDOHLFRGPHFTFSLGSSSLONDMEVSK 720
DB 661 NDNPRLAKDYTGGLFCFHPISAPGSLIFBATDDOHLFRGPHFTFSLGSSSLONDMEVSK 720
QY 721 INGTARLSTRHTDFEBRAYVVLIRINDGSRPPLGIVSLPVTFCSCEGSCRRPAGHOT 780

DB 721 INGTARLSTRHTDFEBRAYVVLIRINDGSRPPLGIVSLPVTFCSCEGSCRRPAGHOT 780
QY 781 GIFTVGAVGIIITLTLVIGIILAVFIRKXKXGNDNSAASEVKPLRS 832
DB 781 GIFTVGAVGIIITLTLVIGIILAVFIRKXKXGNDNSAASEVKPLRS 832

RESULT 4

US-10-025-380-1087

Sequence 1087, Application US/10025380
Publication No. US20020182191A1
GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Stolk, John A.
APPLICANT: Wang, Tongtong
APPLICANT: Jiang, Yagui
APPLICANT: Smith, Carole L.
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Skeiky, Yasir A. W.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick Thomas S.
APPLICANT: Carter, Darrick
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
FILE REFERENCE: 210121.471C14
CURRENT APPLICATION NUMBER: US/10/025,380
NUMBER OF SEQ ID NOS: 1129
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1087
LENGTH: 958
TYPE: PR1
ORGANISM: Homo sapiens
US-10-025-380-1087

Query Match 97.4%; Score 4208; DB 9; Length 958;
Best Local Similarity 99.4%; Pred. No. 5,9e-307;
Matches 813; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 17 LANG-YGQKSGKSGPLKMTFSIYEGQEPSQIIIFQKPNPAVTFELTGETNIFYIER 74
DB 141 LANGSAAQEGKSGKSGPLKMTFSIYEGQEPSQIIIFQKPNPAVTFELTGETNIFYIER 200
QY 75 EGLLYNRALDRETRSTHNOVALDANGIYEGPPIITEVDINDNPTFLOSKEYS 134
DB 201 EGLLYNRALDRETRSTHNOVALDANGIYEGPPIITEVDINDNPTFLOSKEYS 260
QY 135 VRONSRRKPLVYNATDLDPPATNGOLYYQIVOLPMINMYFQINNKTAISLTRE 194
DB 261 VRONSRRKPLVYNATDLDPPATNGOLYYQIVOLPMINMYFQINNKTAISLTRE 320
QY 195 GSGELNPAKNPSYNLVISVDMGQSGENSESDTTSVDIIVTENIMKAPKEVENENSTDP 254
DB 321 GSGELNPAKNPSYNLVISVDMGQSGENSESDTTSVDIIVTENIMKAPKEVENENSTDP 380
QY 255 HPKIKITQVRWNDGAQYSLVDKELRPFPSIDQEBDIYVTOPLDBEEDKAYFVAVAKD 314
DB 381 HPKIKITQVRWNDGAQYSLVDKELRPFPSIDQEBDIYVTOPLDBEEDKAYFVAVAKD 440
QY 315 EYKSPSYPLEIHVKYKNDINDNPTCSPVTVFEVQENERLGNISIGTLTFAHDEBENTAN 374
DB 441 EYKSPSYPLEIHVKYKNDINDNPTCSPVTVFEVQENERLGNISIGTLTFAHDEBENTAN 500
QY 375 SFANRYIVEQTPKLPMDGFLIQTIVAGMLQAKOSLKKODTPQYNLTIEVSDQFTKLCF 434
DB 501 SFANRYIVEQTPKLPMDGFLIQTIVAGMLQAKOSLKKODTPQYNLTIEVSDQFTKLCF 560

QY 435 VOINVINDINQPIFEKSDYGNLTADPTNIGSTLTIOATDADEPFTGSSKILYHIK 494
DB 561 VOINVINDINQPIFEKSDYGNLTADPTNIGSTLTIOATDADEPFTGSSKILYHIK 620
QY 495 DSEGRGLVDTPHTNTGYVLIKKPLDEETAANSIVFKAENPEPLVGVKNASSPAKFT 554
DB 621 DSEGRGLVDTPHTNTGYVLIKKPLDEETAANSIVFKAENPEPLVGVKNASSPAKFT 680
QY 555 LITVDNEAPQFSGHVFQAKVSEDAIGTKVGNVTAADPEGLDISYSLRGDTGWLKIDH 614
DB 681 LITVDNEAPQFSGHVFQAKVSEDAIGTKVGNVTAADPEGLDISYSLRGDTGWLKIDH 740
QY 615 VTGEISVAPLDRAGSPYRVQVAVTEVGSLSVSEFHLIMDVNDNPRILAKDYGL 674
DB 741 VTGEISVAPLDRAGSPYRVQVAVTEVGSLSVSEFHLIMDVNDNPRILAKDYGL 800
QY 675 FFCGPLAPGSLIFEATDDOHLFRGPHFTFSLGSSSLQNDMEVSKINGTHARLSTRHTD 734
DB 801 FFCGPLAPGSLIFEATDDOHLFRGPHFTFSLGSSSLQNDMEVSKINGTHARLSTRHTD 860
QY 735 FEERAYVVLIRINDGRRPLEGIVSLPVTFCSCEVSCCFRPAHQGTGIPVGMAGVILTT 794
DB 861 FEERAYVVLIRINDGRRPLEGIVSLPVTFCSCEVSCCFRPAHQGTGIPVGMAGVILTT 920
QY 795 TLVIGIILAVFIRIKKDKGNVESAGQASEVPLRS 832
DB 921 TLVIGIILAVFIRIKKDKGNVESAGQASEVPLRS 958

RESULT 5

US-09-922-217-1087
Sequence 1087, Application US/09922217
Patent No. US20020076414A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Lodes, Michael J.
APPLICANT: Secretic, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Stolk, John A.
APPLICANT: Wang, Tonglong
APPLICANT: Jiang, Yugu
APPLICANT: Smith, Carole Lynn
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aljun
APPLICANT: Clapper, Jonathan D.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
FILE REFERENCE: 210121.471C13
CURRENT APPLICATION NUMBER: US/09/922.217
NUMBER OF SEQ ID NOS: 1124
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1087
LENGTH: 958
TYPE: PRT
ORGANISM: Homo sapiens
US-09-922-217-1087

Query Match 97.4%; Score 4208; DB 10; Length 958;
Best Local Similarity 99.4%; Pred. No. 5.9e-307;
Matches 813; Conservative 1; Mismatches 2; Indels 2; Gaps 1;
QY 17 LATG--YGOEGKFGSPLKMTFSIYEGQPSQIIFOFKANPAYTFELTGTNDIYVIER 74
DB 141 LASSGAAAGKFGSPLKMTFSIYEGQPSQIIFOFKANPAYTFELTGTNDIYVIER 200
QY 75 EGLLYNRALDRFTRSTNLOVALDANGIIVGEPVPTTIVKQINDNRPFLOSKEYS 134
DB 201 EGLLYNRALDRFTRSTNLOVALDANGIIVGEPVPTTIVKQINDNRPFLOSKEYS 260
QY 135 VRQNSRPGKFLVYNATDLDPAFPGNGLYQVITGLMNNWVFOJNNKTGAISLTRE 194

DB 261 VRQNSRPGKFLVYNATDLDPAFPGNGLYQVITGLMNNWVFOJNNKTGAISLTRE 320
QY 195 GSGEOLNPAKPSNLYIVSKDMGSGENSESDTTSVDIIVTENIMKAPKRVENESTDP 254
DB 321 GSGEOLNPAKPSNLYIVSKDMGSGENSESDTTSVDIIVTENIMKAPKRVENESTDP 380
QY 255 HPKIKTVNRNDGAQYSLVDEKELPRFPESIDQEGDIYVTOPLDREKDAYFVAVAKD 314
DB 381 HPKIKTVNRNDGAQYSLVDEKELPRFPESIDQEGDIYVTOPLDREKDAYFVAVAKD 440
QY 315 EYKPLSLYPELHVKKYKDINDNPTCPSPVYVEVOENERLGNISGTLTAHDEENTAN 374
DB 441 EYKPLSLYPELHVKKYKDINDNPTCPSPVYVEVOENERLGNISGTLTAHDEENTAN 500
QY 375 SFNLYRIVEQTPKLPMDGFLITVYAGMLQALQSLKKQDTPQYNLTIEVSDQFTLCF 434
DB 501 SFNLYRIVEQTPKLPMDGFLITVYAGMLQALQSLKKQDTPQYNLTIEVSDQFTLCF 560
QY 435 VOINVINDINQPIFEKSDYGNLTADPTNIGSTLTIOATDADEPFTGSSKILYHIK 494
DB 561 VOINVINDINQPIFEKSDYGNLTADPTNIGSTLTIOATDADEPFTGSSKILYHIK 620
QY 495 DSEGRGLVDTPHTNTGYVLIKKPLDEETAANSIVFKAENPEPLVGVKNASSPAKFT 554
DB 621 DSEGRGLVDTPHTNTGYVLIKKPLDEETAANSIVFKAENPEPLVGVKNASSPAKFT 680
QY 555 LITVDNEAPQFSGHVFQAKVSEDAIGTKVGNVTAADPEGLDISYSLRGDTGWLKIDH 614
DB 681 LITVDNEAPQFSGHVFQAKVSEDAIGTKVGNVTAADPEGLDISYSLRGDTGWLKIDH 740
QY 615 VTGEISVAPLDRAGSPYRVQVAVTEVGSLSVSEFHLIMDVNDNPRILAKDYGL 674
DB 741 VTGEISVAPLDRAGSPYRVQVAVTEVGSLSVSEFHLIMDVNDNPRILAKDYGL 800
QY 735 FEERAYVVLIRINDGRRPLEGIVSLPVTFCSCEVSCCFRPAHQGTGIPVGMAGVILTT 794
DB 861 FEERAYVVLIRINDGRRPLEGIVSLPVTFCSCEVSCCFRPAHQGTGIPVGMAGVILTT 920
QY 795 TLVIGIILAVFIRIKKDKGNVESAGQASEVPLRS 832
DB 921 TLVIGIILAVFIRIKKDKGNVESAGQASEVPLRS 958

RESULT 6

US-09-833-263-1087
Sequence 1087, Application US/09833263
Patent No. US20020110547A1
GENERAL INFORMATION:
APPLICANT: Wang, Aljun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Stolk, John A.
APPLICANT: Meagher, Madeleine J.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
FILE REFERENCE: 210121.471C12
CURRENT APPLICATION NUMBER: US/09/833.263
NUMBER OF SEQ ID NOS: 1093
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1087
LENGTH: 958
TYPE: PRT
ORGANISM: Homo sapiens
US-09-833-263-1087

Query Match 97.4%; Score 4208; DB 10; Length 958;
Best Local Similarity 99.4%; Pred. No. 5.9e-307;
Matches 813; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

```

QY 17 LATG--YGOBGFSGPLKPMTPFSIYEGEPQIIFQFANPAPVTFELTGTETNIPIYER 74
DB 141 LASGSAAOBCKFGSPLEKPMTPFSIYEGEPQIIFQFANPAPVTFELTGTETNIPIYER 200
QY 75 EGLLYNRALDRETRSTNLOVALDANGIIVEGPVPTIEVKOINDRPFLOSKEYS 134
DB 201 EGLLYNRALDRETRSTNLOVALDANGIIVEGPVPTIEVKOINDRPFLOSKEYS 260
QY 135 VRQNSRPGKPELYNATDLDPATPNQGLYYQIYIQLPMINWYFQINNTGAIISLTRE 194
DB 261 VRQNSRPGKPELYNATDLDPATPNQGLYYQIYIQLPMINWYFQINNTGAIISLTRE 320
QY 195 GSQELNPAKNSYNLVISVKMGQSGSENSFSDTTSVDIIVTENIWKAPKPEWENSTDP 254
DB 321 GSQELNPAKNSYNLVISVKMGQSGSENSFSDTTSVDIIVTENIWKAPKPEWENSTDP 380
QY 255 HPIKITQVRMNDPCAQYSLVDKEKLPFRPFSIDQEGIYVTOPLDREKDAYVYVAVKD 314
DB 381 HPIKITQVRMNDPCAQYSLVDKEKLPFRPFSIDQEGIYVTOPLDREKDAYVYVAVKD 440
QY 315 EYGPRLSYPLEIHKVNDINDNPTCPSPVTVFEVQENRIGNSIGTLTAHDBRENTAN 374
DB 441 EYGPRLSYPLEIHKVNDINDNPTCPSPVTVFEVQENRIGNSIGTLTAHDBRENTAN 500
QY 375 SFLNRYIVEOTPKLPMDGLFIQTYAGMLQAKQSLKKODTPQYNTLIEVSDKDFKTLCF 434
DB 501 SFLNRYIVEOTPKLPMDGLFIQTYAGMLQAKQSLKKODTPQYNTLIEVSDKDFKTLCF 560
QY 435 VOINVIDINDQIPFEKSDYGNLTLAEDTNGSTILLIQAOTDADBPRTGSSKILYHILKG 494
DB 561 VOINVIDINDQIPFEKSDYGNLTLAEDTNGSTILLIQAOTDADBPRTGSSKILYHILKG 620
QY 495 DSEGRGLVDTPDHTNTGYVIIKKPLDEPETAANSIVFEKAPNEPELVGVKNASSPAKFT 554
DB 621 DSEGRGLVDTPDHTNTGYVIIKKPLDEPETAANSIVFEKAPNEPELVGVKNASSPAKFT 680
QY 555 LIYTDVNEAPQFSQHFQAKVSEDAVIGTIVGNVTAKDPEGLDISYSLRGDTRGMLKIDH 614
DB 681 LIYTDVNEAPQFSQHFQAKVSEDAVIGTIVGNVTAKDPEGLDISYSLRGDTRGMLKIDH 740
QY 615 VTGEIFSVAPLDBEAGSPYVQVAVATEVGGSSLSVSSEFHLIMDVNDNPPRLAKDTGTL 674
DB 741 VTGEIFSVAPLDBEAGSPYVQVAVATEVGGSSLSVSSEFHLIMDVNDNPPRLAKDTGTL 800
QY 675 FFCPLSPAPGSLIFEATDDOHLFRGPHFTFSIGSGSLQNDMEVSKINGTHARLSTRHTD 734
DB 801 FFCPLSPAPGSLIFEATDDOHLFRGPHFTFSIGSGSLQNDMEVSKINGTHARLSTRHTD 860
QY 735 FERRAYVVLIRINDGGRPLEGIVSLPVTECSCVBSGCFRPAHQGTGIPVGNVAVGILLT 794
DB 861 FERRAYVVLIRINDGGRPLEGIVSLPVTECSCVBSGCFRPAHQGTGIPVGNVAVGILLT 920
QY 795 TLLVIGIILAVVIRIKKQKGNVESAQASEVYKPLRS 832
DB 921 TLLVIGIILAVVIRIKKQKGNVESAQASEVYKPLRS 958

```

```

RESULT 7
US-10-063-547-98
; Sequence 98, Application US/10063547
; Publication No. US20020182638A1
; GENERAL INFORMATION:
; APPLICANT: Ealon, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Matanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

```

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,547
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 98
; LENGTH: 807
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-547-98

```

```

Query Match 21.0%; Score 908.5; DB 9; Length 807;
Best Local Similarity 30.2%; Pred. No. 1.4e-59;
Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;

```

```

QY 62 LTGE---TDNIVIERE--GLLYNRALDRETRSTNLOVALDANGIIVEGPVPTIEV 116
DB 57 LSGDSGKATEGPAMPDPSGFLVTRALDREQAQVQLQVTLMOQGHVLMGPOVLVHV 116
QY 117 KOINDRPFLOSKEYSVRQNSRPGKPELYNATDLDPATPNQGLYYQIYIQLPMINN 176
DB 117 KOINDRPFLOSKEYSVRQNSRPGKPELYNATDLDPATPNQGLYYQIYIQLPMINN 176
QY 117 KMENDQVPHPSQALRYARLSRGTPGIPFLFELASDRDEGTANSDLRFHILSQAPAPQS 176
DB 117 KMENDQVPHPSQALRYARLSRGTPGIPFLFELASDRDEGTANSDLRFHILSQAPAPQS 176
QY 177 VMYFQINNTGAIISLTREBQSEINPAKNSYNLVISVKMGQSGSENSFSDTTSVDIIVTE 236
DB 177 VMYFQINNTGAIISLTREBQSEINPAKNSYNLVISVKMGQSGSENSFSDTTSVDIIVTE 236
QY 177 PMFQLEPRGALALPKGSTSLDHALBERTYQLLVQVKMGQDA--SGHQATATVEVSIIE 235
DB 177 PMFQLEPRGALALPKGSTSLDHALBERTYQLLVQVKMGQDA--SGHQATATVEVSIIE 235
QY 237 NIWKAPEVWENSTDPHPDIKITQVRMNDPCAQYSLVDKEKLPFRPFSIDQEGDIYVTO 296
DB 237 NIWKAPEVWENSTDPHPDIKITQVRMNDPCAQYSLVDKEKLPFRPFSIDQEGDIYVTO 296
QY 236 STWVSLPRHLLENLVLYPHHMAQVHWSGQDVHYH---ESHPPPEFVNAEGNLVYTR 292
DB 236 STWVSLPRHLLENLVLYPHHMAQVHWSGQDVHYH---ESHPPPEFVNAEGNLVYTR 292
QY 297 PLDREKDAYVYVAVKDEYKPELSEIHKVNDINDNPTCPSPVTVFEVQENRIGL 356
DB 297 PLDREKDAYVYVAVKDEYKPELSEIHKVNDINDNPTCPSPVTVFEVQENRIGL 356
QY 293 ELDRKQAEYLVQRAQNSHGEDYAPLELHVLMENNVNCPDPRDPTVSIPELSPRG 352
DB 293 ELDRKQAEYLVQRAQNSHGEDYAPLELHVLMENNVNCPDPRDPTVSIPELSPRG 352
QY 357 NSIGTLTAHDBRENTANFNLNRYIVEOTPKLPMDG--LFLIQTAGMLQAKQSLKK--OD 414
DB 357 NSIGTLTAHDBRENTANFNLNRYIVEOTPKLPMDG--LFLIQTAGMLQAKQSLKK--OD 414
QY 353 TEVTRLSADADAPGSPNHVYQLSPPEBQVEBRAQVDPSTSSVTLGVLPAAQN 412
DB 353 TEVTRLSADADAPGSPNHVYQLSPPEBQVEBRAQVDPSTSSVTLGVLPAAQN 412
QY 415 TPQYNLTIEVSDK--FKTLCFYQINVIDINDQIPFEKSDYGNLTLAEDTNGSTILLI 472
DB 415 TPQYNLTIEVSDK--FKTLCFYQINVIDINDQIPFEKSDYGNLTLAEDTNGSTILLI 472
QY 413 ILLVLAMDLAGEGFSSTCEVEVAVIDINHAPFITSQIGISLPDVEBGTIVAML 472
DB 413 ILLVLAMDLAGEGFSSTCEVEVAVIDINHAPFITSQIGISLPDVEBGTIVAML 472
QY 473 QATDAD--EFTGSSKIL--YHIKDSSEGLGYDTPDHTNTGYVYIR--KPDEPETAASN 528
DB 473 QATDAD--EFTGSSKIL--YHIKDSSEGLGYDTPDHTNTGYVYIR--KPDEPETAASN 528
QY 473 TAIDADLEP--AFRLMDPAIERGDTGTFGLDWEF--DSGHVRLTLCNLSYEAPASHE 527
DB 473 TAIDADLEP--AFRLMDPAIERGDTGTFGLDWEF--DSGHVRLTLCNLSYEAPASHE 527
QY 529 IYFKAPNEPELVGVKNASSPAKFTLIYTDVNEAPQFSQHFQAKVSEDAVIGTIVGNV 588
DB 529 IYFKAPNEPELVGVKNASSPAKFTLIYTDVNEAPQFSQHFQAKVSEDAVIGTIVGNV 588
QY 528 VVVVVGSAKLV--GPGPGGATATVTVLYVERVMPPKLDQESYEASVPISAPGSLTLTI 586
DB 528 VVVVVGSAKLV--GPGPGGATATVTVLYVERVMPPKLDQESYEASVPISAPGSLTLTI 586
QY 589 TAKDEPGLDISYSLRGDTRGMLKIDHVTGEIFSVAPLD--REAGSPYVQVATEVGGSSL 647
DB 589 TAKDEPGLDISYSLRGDTRGMLKIDHVTGEIFSVAPLD--REAGSPYVQVATEVGGSSL 647
QY 587 QPSDPTISRTLRSLVNDSEGMCLIEKFSGEVHTAQSLOQAQGDFTVTVLVEAQDTA---- 642
DB 587 QPSDPTISRTLRSLVNDSEGMCLIEKFSGEVHTAQSLOQAQGDFTVTVLVEAQDTA---- 642
QY 648 SSVSEFHLIMDVNDNPPRLADYTGILFRCHPLSPAGSLIF--EATDDOHLFRGPHFTS 706
DB 648 SSVSEFHLIMDVNDNPPRLADYTGILFRCHPLSPAGSLIF--EATDDOHLFRGPHFTS 706
QY 643 -----LTLAVPQO-----YICTRQDHGLVSGSPSKDPLDASGHP--YSPT 683
DB 643 -----LTLAVPQO-----YICTRQDHGLVSGSPSKDPLDASGHP--YSPT 683
QY 707 LG--SGSLQNDMEVSKINGTHARLSTRHTDFERRAYVVLIRINDGGRPLEGIVSLPVTEC 765
DB 707 LG--SGSLQNDMEVSKINGTHARLSTRHTDFERRAYVVLIRINDGGRPLEGIVSLPVTEC 765
QY 684 LGPNPTVQGDWRLQTLNLSHAYTLTALHVEERHIIIPVYSHNQ---MMQLLRAVLYC 740
DB 684 LGPNPTVQGDWRLQTLNLSHAYTLTALHVEERHIIIPVYSHNQ---MMQLLRAVLYC 740
QY 766 SC--VBSGCFRPAHQGTGIPVGNVAVGILLTLLVIGIILAVVIRIKKQKGNVESAQA 824
DB 766 SC--VBSGCFRPAHQGTGIPVGNVAVGILLTLLVIGIILAVVIRIKKQKGNVESAQA 824
QY 741 RCNVEGQCKRKVKGMKGMFTKLSAVGILVGLTVAIGIFILIFTHTMWSRKDDPPQADS 800
DB 741 RCNVEGQCKRKVKGMKGMFTKLSAVGILVGLTVAIGIFILIFTHTMWSRKDDPPQADS 800
QY 825 SEVK 828
DB 801 VPLK 804

```

RESULT 8

US-10-063-616-98
 / Sequence 98, Application US/10063616
 / Publication No. US20030013855A1
 / GENERAL INFORMATION:
 / APPLICANT: Eaton, Dan L.
 / APPLICANT: Filvaroff, Ellen
 / APPLICANT: Gerritsen, Mary E.
 / APPLICANT: Goddard, Audrey
 / APPLICANT: Godowski, Paul J.
 / APPLICANT: Grimaldi, Christopher J.
 / APPLICANT: Gurney, Austin L.
 / APPLICANT: Macanabe, Colin K.
 / APPLICANT: Wood, William I.
 / TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 / FILE REFERENCE: P3230R1C1
 / CURRENT APPLICATION NUMBER: US/10/063,616
 / CURRENT FILING DATE: 2002-05-03
 / Prior Application removed - See File Wrapper or Palm
 / NUMBER OF SEQ ID NOS: 170
 / SEQ ID NO 98
 / LENGTH: 807
 / TYPE: PRT
 / ORGANISM: Homo Sapien
 US-10-063-616-98

Query Match 21.0%; Score 908.5; DB 9; Length 807;
 Best Local Similarity 30.2%; Pred. No. 1.4e-59;
 Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;

62 LTGE---TDNIFIERE-GLYYNRALDRETRSTHNLQVLAALDANGIYEGVPTITVEV 116
 57 LSGSGKATEPFPMDPSGFLVTRALDREDAEYQVLEMODGHVLMGPVLAHV 116
 117 KDINDNRPFLQSKYEGSVQNSRPGKPELYVNAATLDDPATNGOLYYQIVQLPMINN 176
 117 KDENDQVPHFSQAIYRARLSRGTGPGIFPLFLASDDEPTANSIDRPHILSQAPQPS 176
 177 VMTFOINKKGAISLTREGSOELNPAKNPSYNIIVSYKMGSGSENSFSDTTSVDIYTE 236
 177 PDMFQLEPRRGLALSPKSGTSLDHALERTYQLLVQVMDGDA-SGHQATATVEVSIIE 235
 237 NIKAPRPEVMEVNSTDPHPKIKITQVRNPDGAQYSLVDKELRRPFSIDQEGDIYVQ 296
 236 STWVSLERPIHLAENLKVLPYHMAQVMSGSDVHYHL---ESHPPGPEVNAEGNLVYTR 292
 297 PLDREKDAVYFVAVAKDEYKPLSYPLEIHVKYKDIINDNPTCPSPVTVFEVOENRIG 356
 293 ELDEAQAQAYLLQVRAQNSHGEDYAAPLELHVLMENDNVPICPRPDPTVSIPELSPPG 352
 357 NSIGTLTAHDEENTANSFLNRYIVEQTPKLPMDG-LFLIQYAGMLQAKOSLKK-QD 414
 353 TEVTRLSAEDADAGSPNSHVYQLLSPEPEDGVEGAFQVDPRTSGSVTLGLVPLRAGON 412
 415 TPQYNLTIEVSDKO--FKTLCFVQIVINDINDQIPFEKSDYGNLTLAEDTNISITLTI 472
 413 ILLVLAMDLAGAGGSSSTCEVEAVATDINDHAPEFTSQIGPISLPEDEVERGTLVAML 472
 473 QATDAD-EPFTGSSKIL-YHIKDSSEGLRGLVDTDPHTNTGYLIK--KPLDFTAAVSN 528
 473 TAIADADLEP---ARLMDFAIERDGTBGTPLDWER--DSGHVRLRLCKNLSYAAASHHE 527
 529 IVFAENDEPLVFGVKTNASSFAKFTLLVTDVNEAPQFSQHVFOAKVSEDAVIGTKVGNV 588
 528 VVVVVVAVAKLV-GPGPGPGATATVTVVERVMPPEPKLDQESYASVYISAPAGSFLTLTI 586
 589 TAKPREGIDISYRGDTRGWLKIDHTYGEIIFSAVPLD-REAGSPYRQVAVATEVGGSSL 647
 587 QPSPPIKTRTFSLVNSSEGWLCEKFSGEVHTAQSLOGAOPGDTTYTLVLAQDTA---- 642
 648 SSVSEFHLIMDVNDNPPRLAKDYGTLFFCPLSLAPGSLIF-EATDDQHLFRGPHFFS 706

Db 643 -----LTLAPVPSQ-----YLCTPRQDHGLIVSGPSKDPDLASGHP-YSFT 663
 Qy 707 LG-SGSLQNDMEVSKINGTHARLSTRHTDPEERAAVVLIRINDGRRPLEGIYSLPVTFC 765
 Db 684 LGSNPTVQDRWRKQTLNGSAVLTLLALHWEVEREHIIPVVSNAQ---MQQLVRYIVC 740
 Qy 766 SC-VEGSCPRPAGHOTGIPTVGNVAGILLTTLVLVIGIILAVERIRIKKQKGNVESAQA 824
 Db 741 RNVGEGQCMRKVRGMKMGWPKLSAVGILVGTILVAGIFILLPTHTWMSRKQPDQPADS 800
 Qy 825 SEVX 828
 Db 801 VPLK 804

RESULT 9

US-10-063-502-98
 / Sequence 98, Application US/10063502
 / Publication No. US20030023042A1
 / GENERAL INFORMATION:
 / APPLICANT: Eaton, Dan L.
 / APPLICANT: Filvaroff, Ellen
 / APPLICANT: Gerritsen, Mary E.
 / APPLICANT: Goddard, Audrey
 / APPLICANT: Godowski, Paul J.
 / APPLICANT: Grimaldi, Christopher J.
 / APPLICANT: Gurney, Austin L.
 / APPLICANT: Macanabe, Colin K.
 / APPLICANT: Wood, William I.
 / TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 / FILE REFERENCE: P3230R1C1
 / CURRENT APPLICATION NUMBER: US/10/063,502
 / CURRENT FILING DATE: 2002-05-01
 / Prior Application removed - See File Wrapper or Palm
 / NUMBER OF SEQ ID NOS: 170
 / SEQ ID NO 98
 / LENGTH: 807
 / TYPE: PRT
 / ORGANISM: Homo Sapien
 US-10-063-502-98

Query Match 21.0%; Score 908.5; DB 9; Length 807;
 Best Local Similarity 30.2%; Pred. No. 1.4e-59;
 Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;

62 LTGE---TDNIFIERE-GLYYNRALDRETRSTHNLQVLAALDANGIYEGVPTITVEV 116
 57 LSGSGKATEPFPMDPSGFLVTRALDREDAEYQVLEMODGHVLMGPVLAHV 116
 117 KDINDNRPFLQSKYEGSVQNSRPGKPELYVNAATLDDPATNGOLYYQIVQLPMINN 176
 117 KDENDQVPHFSQAIYRARLSRGTGPGIFPLFLASDDEPTANSIDRPHILSQAPQPS 176
 177 VMTFOINKKGAISLTREGSOELNPAKNPSYNIIVSYKMGSGSENSFSDTTSVDIYTE 236
 177 PDMFQLEPRRGLALSPKSGTSLDHALERTYQLLVQVMDGDA-SGHQATATVEVSIIE 235
 237 NIKAPRPEVMEVNSTDPHPKIKITQVRNPDGAQYSLVDKELRRPFSIDQEGDIYVQ 296
 236 STWVSLERPIHLAENLKVLPYHMAQVMSGSDVHYHL---ESHPPGPEVNAEGNLVYTR 292
 297 PLDREKDAVYFVAVAKDEYKPLSYPLEIHVKYKDIINDNPTCPSPVTVFEVOENRIG 356
 293 ELDEAQAQAYLLQVRAQNSHGEDYAAPLELHVLMENDNVPICPRPDPTVSIPELSPPG 352
 357 NSIGTLTAHDEENTANSFLNRYIVEQTPKLPMDG-LFLIQYAGMLQAKOSLKK-QD 414
 353 TEVTRLSAEDADAGSPNSHVYQLLSPEPEDGVEGAFQVDPRTSGSVTLGLVPLRAGON 412
 415 TPQYNLTIEVSDKO--FKTLCFVQIVINDINDQIPFEKSDYGNLTLAEDTNISITLTI 472
 413 ILLVLAMDLAGAGGSSSTCEVEAVATDINDHAPEFTSQIGPISLPEDEVERGTLVAML 472

PRIOR FILING DATE:	1998-09-25
PRIOR APPLICATION NUMBER:	60/101916
PRIOR FILING DATE:	1998-09-24
PRIOR APPLICATION NUMBER:	60/101922
PRIOR FILING DATE:	1998-09-24
PRIOR APPLICATION NUMBER:	60/106178
PRIOR FILING DATE:	1998-10-28
PRIOR APPLICATION NUMBER:	60/106248
PRIOR FILING DATE:	1998-10-29
PRIOR APPLICATION NUMBER:	60/106464
PRIOR FILING DATE:	1998-10-30
PRIOR APPLICATION NUMBER:	60/106905
PRIOR FILING DATE:	1998-11-03
PRIOR APPLICATION NUMBER:	60/108787
PRIOR FILING DATE:	1998-11-17
PRIOR APPLICATION NUMBER:	60/108801
PRIOR FILING DATE:	1998-11-17
PRIOR APPLICATION NUMBER:	60/108849
PRIOR FILING DATE:	1998-11-18
PRIOR APPLICATION NUMBER:	60/112422
PRIOR FILING DATE:	1998-12-15
PRIOR APPLICATION NUMBER:	60/113296
PRIOR FILING DATE:	1998-12-22
PRIOR APPLICATION NUMBER:	60/113605
PRIOR FILING DATE:	1998-12-23
PRIOR APPLICATION NUMBER:	60/113621
PRIOR FILING DATE:	1998-12-23
PRIOR APPLICATION NUMBER:	60/115558
PRIOR FILING DATE:	1999-01-12
PRIOR APPLICATION NUMBER:	60/115565
PRIOR FILING DATE:	1999-01-12
PRIOR APPLICATION NUMBER:	60/115753
PRIOR FILING DATE:	1999-01-12
PRIOR APPLICATION NUMBER:	60/119549
PRIOR FILING DATE:	1999-02-10
PRIOR APPLICATION NUMBER:	60/123618
PRIOR FILING DATE:	1999-03-10
PRIOR APPLICATION NUMBER:	60/125259
PRIOR FILING DATE:	1999-03-19
PRIOR APPLICATION NUMBER:	60/125775
PRIOR FILING DATE:	1999-03-23
PRIOR APPLICATION NUMBER:	60/126773
PRIOR FILING DATE:	1999-03-29
PRIOR APPLICATION NUMBER:	60/127887
PRIOR FILING DATE:	1999-04-05
PRIOR APPLICATION NUMBER:	60/130232
PRIOR FILING DATE:	1999-04-21
PRIOR APPLICATION NUMBER:	60/131022
PRIOR FILING DATE:	1999-04-26
PRIOR APPLICATION NUMBER:	60/131270
PRIOR FILING DATE:	1999-04-27
PRIOR APPLICATION NUMBER:	60/131293
PRIOR FILING DATE:	1999-04-27
PRIOR APPLICATION NUMBER:	60/131445
PRIOR FILING DATE:	1999-04-28
PRIOR APPLICATION NUMBER:	60/134287
PRIOR FILING DATE:	1999-05-14
PRIOR APPLICATION NUMBER:	60/140650
PRIOR FILING DATE:	1999-06-22
PRIOR APPLICATION NUMBER:	60/140723
PRIOR FILING DATE:	1999-06-22
PRIOR APPLICATION NUMBER:	60/141037
PRIOR FILING DATE:	1999-06-23
PRIOR APPLICATION NUMBER:	60/144758
PRIOR FILING DATE:	1999-07-20
PRIOR APPLICATION NUMBER:	60/145688
PRIOR FILING DATE:	1999-07-26
PRIOR APPLICATION NUMBER:	60/146222
PRIOR FILING DATE:	1999-07-28
PRIOR APPLICATION NUMBER:	60/146963
PRIOR FILING DATE:	1999-08-03
PRIOR APPLICATION NUMBER:	60/149320
PRIOR FILING DATE:	1999-08-17

[illegible]

RESULT 11
US-10-230-163-188
Sequence 188: Application US/10230163
Publication No. US2003036635A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Desnoyers, Luc
APPLICANT: Gerritsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe F.
APPLICANT: Watanabe, Colin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3510PIC96
CURRENT FILING DATE: 2002-08-28
CURRENT FILING DATE: 2002-08-28
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/089552
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089558
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089905
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090691
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

PRIOR APPLICATION NUMBER: 60/095302
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095318
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095916
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/096146
PRIOR FILING DATE: 1998-08-11
PRIOR APPLICATION NUMBER: 60/096791
PRIOR FILING DATE: 1998-08-17
PRIOR APPLICATION NUMBER: 60/097986
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/098544
PRIOR FILING DATE: 1998-08-31
PRIOR APPLICATION NUMBER: 60/099596
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100038
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106178
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/106248
PRIOR FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: 60/106464
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 60/106905
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/108787
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108801
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108849
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115558


```

PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/125259
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 60/125775
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/126773
PRIOR FILING DATE: 1999-03-29
PRIOR APPLICATION NUMBER: 60/127887
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/130232
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131022
PRIOR FILING DATE: 1999-04-26
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131445
PRIOR FILING DATE: 1999-04-28
PRIOR APPLICATION NUMBER: 60/134287
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/140650
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/140723
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146963
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/149320
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149638
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/151733
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/164418
PRIOR FILING DATE: 1999-11-09
PRIOR APPLICATION NUMBER: 60/166361
PRIOR FILING DATE: 1999-11-16
PRIOR APPLICATION NUMBER: 60/169445
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169835
Query Match 21.0%; Score 908.5; DB 9; Length 807;
Best Local Similarity 30.2%; Pred. No. 1.4e-59;
Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;
QY 62 LTGE---TDNIEVERB-GLLYNRAIDRETRSTHNLQVAALDANGIIEGPIITIEV 116
DB 57 LSGSGKATEGPFAMDPDSGLTALDRREGEQAEYQLQVLEMQDGHVWGPOPIVAVH 116
QY 117 KQINDNPTLOSTYEGSVRNSRGKPELYVNTDIDDPATPGQLYVOIVLQPMNN 176
DB 117 KDENDQVPHFSQALYRRLSRGTBGLPFLFLASDRPEPTANSDLRFHLSQAPOPS 176
QY 177 VMYOINNKGTALSLTREBSQELNPAKNPSYNLVYSVKMGQSGNSFSDTTSVDIYTE 236

```

```

DB 177 PDMFQLEPRIGALALSPKGSSTSLDHALERTYQLLVQVKMDGQA-SGHQATATVEVSILE 235
QY 237 NIMKAPKPVEMWENSTDPHPKIKTOYRWMDPGAOYSLVDKELRPFPSIDOGDIYVTO 296
DB 236 STWVSLPETHLENLKVLTPHHVAQVHMSGGVHNL---ESHPPEPFEVNAEGNLYVTR 292
QY 297 PLDBEKDAVYFAVAKDEYKPLSYBLEIHKVKDINDNPPTCPSPVTFVEQENERIG 356
DB 293 ELDRENAQAEYLLQVRQNSHGEDYANAPLEHLVLMENDNVPICPRPDPTVSIPELSPG 352
QY 357 NSIGTLTAHDREENTANSPFNRYIEQTPKPMDC-LPIQTYAGMLQAKSLK-OD 414
DB 353 TEVTRLASBDADAPGNSHVYQLSPBEBDEVERAQQVPTSGSVTLGVPLPAGON 412
QY 415 TPQVNLTVESDOD--FKTLCPQVQVINDOIPFEKSDVGNLTAEPTNGSTILT 472
DB 413 ILLVLAMDLAGEGFSSTCEVEAVTINDHAFITSOIGPISLPEDVEGTIVML 472
QY 473 QATDAD-EPFTGSSKIL-YHIKDSGRIGVDDPHNTGVYIK--KPLDETAASN 528
DB 473 TAIADALEP---AFRLMDFAIERGDTGTFGLDMEB--DSGHVRLRLCKNLSEEAAPSH 527
QY 529 IVKAKENPEPLVFGVKYNNASFPKFTLYTDVNEAQFGQHVFOQAYSEDAVIGTVANV 588
DB 528 VVVVQSVAKLV-GPGPGGATATVTLVERVWPPKLOQESYEASVPIAPAGSPLT 586
QY 589 TAKDPREGDISYSLRDGTGWLKIDHVTGEIFSAVPLD-REAGSPRYQVAVTEVGSSL 647
DB 587 QPSDPISTRFSLVNDSSBMLCTEKFSGEHTAQSLOQAQPDITVTVLEADPTA---- 642
QY 648 SSVSEFHLILMDVNDNPPRLAKDYGLFCHPLSABGSLF-BATDDQHLFRGPHFTFS 706
DB 643 -----LTLAPVPSQ-----YLCPTPQDHGLVSGPSKDPDLASGCP-YSFT 683
QY 707 LG-SGSLQNDWEVSKINGTHARLSTHTDPEEPAVYVLRINDGSGPPLLEGISLPYTC 765
DB 684 LGNPTVQDRWRLQTLNGSHAYLTLALHWVEPREHIIPVVSHNAQ---MMQLLVRIYVC 740
QY 766 SC-VEGSCRRPAGHQIGPTVGAAGVGLTTLVLVIGIILAVFIRIKDKDNVESQA 824
DB 741 RCNVEGQCKKVKRMKGMPTKLSAVGLVGTILVAILFILITHTMTMSKKDPDPADS 800
QY 825 SEVK 828
DB 801 VPLK 804
RESULT 12
US-10-006-856A-229
; Sequence 229, Application US/10006856A
; Publication No. US20030044841A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Baton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gueney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C14
; CURRENT APPLICATION NUMBER: US/10/006,856A
; NUMBER OF SEQ ID NOS: 477
; Prior Application removed - See File Wrapper or Palm

```



```

Db 413 ILLVLAMDLAAGEGSSSTCEVEAVNTDINDHAPETISQIGPISLPEDEPGLVAML 472
Qy 473 QATDAD-EPTTSSKLL-YHIIKDSGRLGVDTDPHTNTGYIIR--KPLDETAASN 528
Db 473 TAIADALEP---AFRLMDFAIERGDTGTEGLDMEP--DSGHVRLRLCKNLSTEAAPSHS 527
Qy 529 IVFKAENPEPLVGVKNASSPAKFTLLIVTDNEAPQSFQHVQAKVSEDAVIGTKVGN 588
Db 528 VVVVVSQVAKLV-GPGPGPATATVTVLVERVMPPLKDESEASVPIAPAGSFLTTI 586
Qy 589 TAKDPGLDISYSLRQDTRGMLKIDHTGEIIFSVAPLD-REAGSPYRVQVATEVGGSSL 647
Db 587 QPDPDISRTLRFSLVNDSEGLCIKESGCVHTAQSLOGAQPQDGTIVLVEADDTA---- 642
Qy 648 SSVSEFLLIMDVNDNPPRLAKDYGTLFCHPLSAPGSLIF-EATDDQHLFRGPHRTS 706
Db 643 -----LTLAPVPSQ-----YLCTPRQDHGLIVSGPSKPDLSAGHP-YSFT 683
Qy 707 LG-SGSLONDMESKINGTHARLSTRHTDPEERAYVVLIRINDGRRPLEGISLPTVFC 765
Db 684 LGPPVYQRMKRLQTLNGSHAVTLTALHWVEREHHIIPVVASHNAQ---MMQLLVRIYVC 740
Qy 766 SC-VEGSCFRPAGHQTGIPVGNVAVGILLTLLVIGIILAVFLIRIKDKGKNVESAOA 824
Db 741 RCNVEGQCMKRVGMKGMPTKLSAVGILVGTVAIGIFILLIFTHMTMSRKDPDQPADS 800
Qy 825 SEVK 828
Db 801 VPLK 804

RESULT 14
US-10-230-338-188
; Sequence 188, Application US/10230338
; Publication No. US20030044934A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerdtisen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACTS ENCODING THE SAME
; FILE REFERENCE: P3530P1C92
; CURRENT APPLICATION NUMBER: US/10/230.338
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.

```

```

; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 188
; LENGTH: 807
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-230-338-188

Query Match      21.0%; Score 908.5; DB 9; Length 807;
Best Local Similarity 30.2%; Pred. No. 1.4e-59;
Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;

Qy 62 LTGE-----TDNIFVIERE-GLLYNRAALDRETRSTNHLQVAALDANGIIVEGVPIITIEV 116
Db 57 LSGDSGKATGEPFAMPDPSGFLVTRALDREBOAEYQLQVTEMOGHVLMGQPVYLVHY 116
Qy 117 KDINDRPFPLQSGYKSGYRONSRRPKFPLVYNAATLDDPARPNGLYYQIVQLMINN 176
Db 117 KQENDQVPHPSQAIYVARLSRGRPGIPPLFLASDRDEPTANSDLRPHILISQADAPDS 176
Qy 177 VVWFQINNKTGAISLTRREGSOELNPAKNPSYMLVIVKMGQSGSENSFSDTTSVDIIVTE 236
Db 177 PDMFQLEPRLGALALSPKSTSLDHALERTYQLLVQVKNMGDQA-SGHQATATVERISIE 235
Qy 237 NIWKAEPVEMWENSTDPPIKITQYRWNDPGAQYSLVDKELPRFPFSIDQEGDIYVTO 296
Db 236 STWVSLPEIHLAENLKVLYPHHMAQVHMGGDVHYHL---ESHPPGFVEVNAEGNLVYTR 292
Qy 297 PLDREKDAVYFVAVAKDEYKPLSTPLFIHYVKQINDNPPTCPSPVTVFVEQEMERIG 356
Db 293 ELDRQAQAEYLLQVRAQNSHGEDYAPLEHLVLMDENNVPICPRPDPVTVISIPELSPPG 352
Qy 357 NSIGTLTAHDREENTANSFLNRIYEQTPKLPMDG-LFLICTYAGMLQAKOSLKK-OP 414
Db 353 TEYTRLSAEDADAPGSPNSHVYQLSPREBQVEGRAQVQDPTSSVTLGVPLAAGQN 412
Qy 415 TPQYNLTIEVSDK--FKTLCPVQVINDINDQIPFEKSDYGNLTIAEDTNGITTLTI 472
Db 413 ILLVLAMDLAAGEGSSSTCEVEAVNTDINDHAPETISQIGPISLPEDEPGLVAML 472
Qy 473 QATDAD-EPTTSSKLL-YHIIKDSGRLGVDTDPHTNTGYIIR--KPLDETAASN 528
Db 473 TAIADALEP---AFRLMDFAIERGDTGTEGLDMEP--DSGHVRLRLCKNLSTEAAPSHS 527
Qy 529 IVFKAENPEPLVGVKNASSPAKFTLLIVTDNEAPQSFQHVQAKVSEDAVIGTKVGN 588
Db 528 VVVVVSQVAKLV-GPGPGPATATVTVLVERVMPPLKDESEASVPIAPAGSFLTTI 586
Qy 589 TAKDPGLDISYSLRQDTRGMLKIDHTGEIIFSVAPLD-REAGSPYRVQVATEVGGSSL 647
Db 587 QPDPDISRTLRFSLVNDSEGLCIKESGCVHTAQSLOGAQPQDGTIVLVEADDTA---- 642
Qy 648 SSVSEFLLIMDVNDNPPRLAKDYGTLFCHPLSAPGSLIF-EATDDQHLFRGPHRTS 706
Db 643 -----LTLAPVPSQ-----YLCTPRQDHGLIVSGPSKPDLSAGHP-YSFT 683
Qy 707 LG-SGSLONDMESKINGTHARLSTRHTDPEERAYVVLIRINDGRRPLEGISLPTVFC 765
Db 684 LGPPVYQRMKRLQTLNGSHAVTLTALHWVEREHHIIPVVASHNAQ---MMQLLVRIYVC 740
Qy 766 SC-VEGSCFRPAGHQTGIPVGNVAVGILLTLLVIGIILAVFLIRIKDKGKNVESAOA 824
Db 741 RCNVEGQCMKRVGMKGMPTKLSAVGILVGTVAIGIFILLIFTHMTMSRKDPDQPADS 800
Qy 825 SEVK 828
Db 801 VPLK 804

RESULT 15
US-10-063-518-98
; Sequence 98, Application US/10063518
; Publication No. US200300449735A1
; GENERAL INFORMATION:

```

APPLICANT: Eaton, Dan L.
APPLICANT: Filvarcoff, Ellen
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Goddard, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gueney, Austin L.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3230R1C1
CURRENT APPLICATION NUMBER: US/10/063,518
CURRENT FILING DATE: 2002-05-01
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 98
LENGTH: 807
TYPE: PRT
ORGANISM: Homo Sapien
US-10-063-518-98

Query Match 21.0%; Score 908.5; DB 9; Length 807;
Best Local Similarity 30.2%; Pred. No. 1.4e-59;
Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;
QY 62 LTGE---TNIPIERE-GILYRNALDRETRSTHNLQYAAALDANGIIVEGPVITIEV 116
DB 57 LSGDSGKATGPRFMDSDGFLVTRALDREQAQYQQLTLEMQDGHVLMGPOPLVHV 116
QY 117 KDINDNPTFLQSKTEGSSVQRNSRPGKFLVYVATDLDPPATNGQLYQIVIQIPMINN 176
DB 117 KDEPDQVPHFSQAIYRRLSRGTGPIPLFLASDRDEPTANSIDRPHILSQAPQPS 176
QY 177 VMYFQINNKGAISLTREGSQELNPAKPSNIVISKMDGGGSENSFSDTSDIIVTE 236
DB 177 PMFQLEPRIGALALSPKSGTSLDHALERTYQLLVQKMGDOA-SGQATATVEVSIIE 235
QY 237 NIMKAPKPEVMEVNSTDPHPKIKITQVWMDPGAQYSLVDEKELPRPFPSIDQEGDIYVTO 296
DB 236 STWVSEPIHAEHLKLYRHHMAQVWMSGDVHYHL--ESHPPGFVNAEGNLYVTR 292
QY 297 PLDREKDAYFYAAVADKXKPLSYPLKIHVKVKDINDNPTCPSPVTFEVOENERLG 356
DB 293 ELDRQAQAEYLLQVRAQNSHGEDVYAPLHLVMDENDVPICPPRDPTVSIPELSPPG 352
QY 357 NSIGTLTAHRDEENTANSFLNRIYEQPKLPMDG-LFLIQYAGMLQAKOSLKK-QD 414
DB 353 TEVTRLSAEDADAPGSPNSHVYQLSPBEDGVEGRAFOVDPPTSGSVTLGVLPFRAGQN 412
QY 415 TPQYNLTIEVSDKD--FKTLCFQVQVINDINDQPIFEKSDYGNLTLAEDTNIGSTILTI 472
DB 413 ILLVLTAMDLAGAEGFSSICEVEAVTINDHAPERTISQIGFISLPEDVEPRTLVAML 472
QY 473 QATDAD-EPPTGSSKIL-YHIKDGSEGRGLVDTDPHTNTGYVLIK--KPLDEPETAASN 528
DB 473 TALDADLEP--AFRLMDFAIERGDTGTGFLDWER--DSGHVRLRLCKNLSYEAPSHSHE 527
QY 529 IVKKAENPEPLVGVKTNASSFAKFTLIVTDNEAPQFSQHVFOAKYSEDAVIGTKVGNV 588
DB 528 VVVVVQSVAKLV-GPGPGATATVTVLVERVMPPLDQESYEASVPISAPAGSFLTLTI 586
QY 589 TAKDPEDGLDISYLRGRTGMLKIDHTVGEIIFSAPLD-REAGSPYRVQVAVTEVGGSSL 647
DB 587 QPSDPTISRTLRFSLVNDSBGLCTEKFSGEVHTAQSLQGAQPGDTTVLVEADDTA---- 642
QY 648 SSVSEPHLIIIMDVNDNPPRLAKDYTGFLFCHPLSAPGSLIF-EATDDQHLFRGPHFTFS 706
DB 643 -----LTLAPVPSQ-----YLCPTPRQDHGLIVSPSKDPLASGHGP--YSFT 683
QY 707 LG-SGSLQNDWESKINGTHARLSTRITDPEERAYVVLIRINDGGRPPLLEGIVSLPPTFC 765
DB 684 LGNPTVQQRDMRLQTLINGSHAYLTLALHWVEPREHIIIPVVVSHNAQ---NMQLLVRYIVC 740

QY 766 SC-VEGSCFRPAGHQGTGPTVGNVAGILTTLLVIGIILAVFIRIKDKGKNVESQAQ 824
DB 741 RCNVEGQCMKRYGRMKGMPTKLSAVGILVGTVAIGIFLLITHTMTMSRKKOPDGPADS 800
QY 825 SEVK 828
DB 801 VPLK 804

Search completed: June 22, 2003, 01:10:50
Job time : 66 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2003, 01:10:55 ; Search time 274 Seconds
(without alignments)
4455.837 Million cell updates/sec

Title: US-10-025-380-1081
Perfect score: 4321
Sequence: 1 MIQAMHSLCLMLYATG.....DKGKNVSAQASEVPLRS 832

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1042519 seqs, 733713590 residues
Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xid
-O/cgmr2.1/USPTO.spool/US10025380/rnatc.17062003.173329.24718/app.query.fasta.1.1031
-DB=Published.Applications.NA -QFMT=fastap -SUFFIX=rmpb -MINMATCH=0.1
-LOOPCD=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=bl0sum62
-TRANS=human40.csl -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10025380 @CGN 1.1.166 @rnatc.17062003.173329.24718
-NCPU=6 -ICPU=3 -NO MAP -LARGESQURRY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgmr2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgmr2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgmr2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgmr2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgmr2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgmr2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgmr2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgmr2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgmr2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgmr2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgmr2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgmr2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgmr2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgmr2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4321	100.0	3345	9	US-10-025-380-1076
2	4321	100.0	3345	10	US-09-922-217-1076
3	4321	100.0	3345	10	US-09-962-436-302
4	4321	100.0	3345	10	US-09-833-263-1076

5	4321	100.0	3345	10	US-09-880-107-3288	Sequence 3288, Ap
6	4303	99.6	3654	10	US-09-962-832-118	Sequence 118, Ap
7	4303	99.6	3654	10	US-09-880-107-3807	Sequence 3807, Ap
8	4208	97.4	2877	9	US-10-025-380-1086	Sequence 1086, Ap
9	4208	97.4	2877	10	US-09-922-217-1086	Sequence 1086, Ap
10	4208	97.4	2877	10	US-09-833-263-1086	Sequence 1086, Ap
11	908.5	21.0	2848	9	US-10-063-547-97	Sequence 97, Appl
12	908.5	21.0	2848	9	US-10-063-516-97	Sequence 97, Appl
13	908.5	21.0	2848	9	US-10-063-502-97	Sequence 97, Appl
14	908.5	21.0	2848	9	US-10-227-884-187	Sequence 187, Appl
15	908.5	21.0	2848	9	US-10-230-163-187	Sequence 187, Appl
16	908.5	21.0	2848	9	US-10-006-856-228	Sequence 228, Appl
17	908.5	21.0	2848	9	US-10-218-631-187	Sequence 187, Appl
18	908.5	21.0	2848	9	US-10-230-138-187	Sequence 187, Appl
19	908.5	21.0	2848	9	US-10-063-118-97	Sequence 97, Appl
20	908.5	21.0	2848	9	US-10-063-598-97	Sequence 97, Appl
21	908.5	21.0	2848	9	US-10-227-693-97	Sequence 97, Appl
22	908.5	21.0	2848	9	US-10-230-614-187	Sequence 187, Appl
23	908.5	21.0	2848	9	US-10-006-818-228	Sequence 228, Appl
24	908.5	21.0	2848	9	US-10-015-393-228	Sequence 228, Appl
25	908.5	21.0	2848	9	US-10-063-567-97	Sequence 97, Appl
26	908.5	21.0	2848	9	US-10-216-159-187	Sequence 187, Appl
27	908.5	21.0	2848	9	US-09-946-374-228	Sequence 228, Appl
28	908.5	21.0	2848	9	US-10-012-121-228	Sequence 228, Appl
29	908.5	21.0	2848	9	US-10-015-869-228	Sequence 228, Appl
30	908.5	21.0	2848	9	US-10-218-849-187	Sequence 187, Appl
31	908.5	21.0	2848	9	US-10-227-873-187	Sequence 187, Appl
32	908.5	21.0	2848	9	US-10-227-883-187	Sequence 187, Appl
33	908.5	21.0	2848	9	US-10-063-599-97	Sequence 97, Appl
34	908.5	21.0	2848	9	US-10-219-076-187	Sequence 187, Appl
35	908.5	21.0	2848	9	US-10-230-434-187	Sequence 187, Appl
36	908.5	21.0	2848	9	US-10-006-116-228	Sequence 228, Appl
37	908.5	21.0	2848	9	US-10-017-527-228	Sequence 228, Appl
38	908.5	21.0	2848	9	US-10-013-913-228	Sequence 228, Appl
39	908.5	21.0	2848	9	US-10-017-527-228	Sequence 228, Appl
40	908.5	21.0	2848	9	US-10-063-595-97	Sequence 97, Appl
41	908.5	21.0	2848	9	US-10-219-003-187	Sequence 187, Appl
42	908.5	21.0	2848	9	US-10-219-075-187	Sequence 187, Appl
43	908.5	21.0	2848	9	US-10-219-464-187	Sequence 187, Appl
44	908.5	21.0	2848	9	US-10-219-466-187	Sequence 187, Appl
45	908.5	21.0	2848	9	US-10-219-479-187	Sequence 187, Appl

ALIGNMENTS

RESULT 1
US-10-025-380-1076
Sequence 1076, Application US/10025380
Publication No. US20020182191A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Lodges, Michael J.
APPLICANT: Secretist, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Stolk, John A.
APPLICANT: Wang, Tongtong
APPLICANT: Jiang, Yuguu
APPLICANT: Smith, Carole L.
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Skeiky, Yaeir A. W.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick Thomas S.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
FILE REFERENCE: 210121.471c14
CURRENT APPLICATION NUMBER: US/10/025,380
CURRENT FILING DATE: 2001-12-19
NUMBER OF SEQ ID NOS: 1129
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1076
 LENGTH: 3345
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-025-380-1076

Alignment Scores:

Pred. No.:	0	Length:	3345
Score:	4321.00	Matches:	832
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-025-380-1081 (1-832) x US-10-025-380-1076 (1-3345)

Qy 1 MetIleuGlnAlaHisLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
 Db 88 ATGATACCTTCAGGCGCCATCTTCACCTCGTGTCTTCTTATGCTTTATTTGGCAACTGGA 147
 Qy 21 TyrGlyGlnGlyGlyPheSerGlyProLeuYsProMetThrPheSerIleTyrGlu 40
 Db 148 TATGCGCAAGAGGGGAAATTAGTGAAGCCCTGAAACCCATGACATTTCTATTATGA 207
 Qy 41 GlyGlnGlyProSerGlnIleIlePheGlnPheLysAlaAsnProAlaValThrPhe 60
 Db 208 GGGCAAGAACCGAGTCATATTATTCAGTTTAAAGCCATCTCTCTGCTGACTTTT 267
 Qy 61 GluLeuThrGlyGlyLysThrAspAsnIlePheValIleGluArgGlyGlyLeuLeuTyr 80
 Db 268 GAACTTAAGTGGGAGAGACAGACATATTTGTGATAGACGGGAGGACTTCTGTATTAAC 327
 Qy 81 AsnArgAlaLeuAspArgGlyThrArgSerThrHisAsnLeuGlnValAlaLeuAsp 100
 Db 328 AACGAGCGCTTGACAGGGAACAAGATCTACTCAATCTCCAGTTTGCAGCCCTGAC 387
 Qy 101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValLysAspIleAsn 120
 Db 388 GCTATGGAATTTATGAGGAGGATCCATCCCTTACCATGAAAGTGAAGACATCAAC 447
 Qy 121 AspAsnArgProThrPheLeuGlnSerLysTyrGlyGlySerValArgGlnAsnSerArg 140
 Db 448 GACATGAGACCCAGCTTTCTCCAGTCAAAAGTACGAGGCTCAGTAAGCAGAACTTCGC 507
 Qy 141 ProGlyLysProPheLeuTyrValAsnAlaThrAspLeuAspProAlaThrProAsn 160
 Db 508 CCAGAAAGCCCTTCTGTATGATGCACAGACCTGATGATCCGGCCACCTCCCAT 567
 Qy 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
 Db 568 GGGCAGCTTTATCCAGATTGTCTCCAGCTTCCCATGATCAACATGTCTATCTTT 627
 Qy 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGlyGlySerGlnGluLeuAsn 200
 Db 628 CAGATCAACAACAAGGAGGAGCATCTCTTACCCGAGAGGAGTCAAGAAATTGAT 687
 Qy 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlySer 220
 Db 688 CTGCTTAAGATCTTCTTATATCTGTGATCTCAGTGAAGGACATGGAGGGCCAGAT 747
 Qy 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValIleThrGluAsnIleTyrLys 240
 Db 748 GAGATTCCTTCAGTATCAACATCTGTGATATCATGACAGAGAAATTTTGGAA 807
 Qy 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
 Db 808 GCACCAAAACCTGTGAGATGAGTGAAGAACTCACTGATCTCAACCCCATCAAAATCACT 867
 Qy 261 GlnValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGluLysLeuPro 280
 Db 868 CAGGTGCGGTGAATGATCCGGTGCACAATATTTCTTATGTTGACAAAGAAAGCTGCA 927
 Qy 281 ArgPheProPheSerIleAspGlnGlyAspIleTyrValThrGlnProLeuAspArg 300

Db 928 AGATCCCATTTTCAATTGACGAGAGGATATTACGTCACTCAGCCCTTGACCGCA 987
 Qy 301 GluGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGlnTyrGlyLysProLeu 320
 Db 988 GAAAGAAAGATGCAATATGTTTTTTTATGCAAGTTGCAAGAGATGATGCAAGAAACCACTT 1047
 Qy 321 SerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProThrCys 340
 Db 1048 TCATATCCGCTGAAATTTCTATGTAAGTTAAAGATTATATATATATCACTTACATG 1107
 Qy 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 Db 1108 CCGTCAACAGTAAACCGATTTGAGTCCAGAGAAATGAACGACTGGTAACTATCGGG 1167
 Qy 361 ThrLeuThrAlaHisAspArgAspGlyLysAsnThrAlaAsnSerPheLeuAsnTyrArg 380
 Db 1168 ACCCTTACGTGACATACAGAGGATGAAGAAATACGCAACGTTTCTTAACTACAGG 1227
 Qy 381 IleValGlnGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 Db 1228 ATTGTGAGCAAACTCCCAACTTCCCATGATGACTCTTCTTATCCAACTATGCT 1287
 Qy 401 GlyMetLeuGlnLeuValLysGlnSerLeuLysGlyAspThrProGlnTyrAsnLeu 420
 Db 1288 GAATGTTCAGTTAGTAAACAGTCTTGAAAGAACAGATACCTCTCAGTCAACTTA 1347
 Qy 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
 Db 1348 ACGATAGAGGTGTCTGACAAAGATTTCAAGCCCTTGTGTGCAAACTAGTATT 1407
 Qy 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAla 460
 Db 1408 GATATCATGATAGATGCCCATCTTTGAAAATTCAGATTATGAAACCTGACTTCTGCT 1467
 Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db 1468 GAAGACAAACATTTGGGTCCACCATCTTAACCATCAGGCCACTGATGATGAGCCA 1527
 Qy 481 PheThrGlySerSerLysIleLeuTyrHisIleIleLysGlyAspSerGlyLysArgLeu 500
 Db 1528 TTTTACGGGAGTCTTAAATTTCTGTATCATATATTAAGGAAACAGTGAAGGAGCGCTG 1587
 Qy 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
 Db 1588 GGGGTGACACAGATCCCATCAACACCGGATATGTCTAATTTAAAGCCTCTTGAT 1647
 Qy 521 PheGluThrAlaAlaLysSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 Db 1648 TTTGAAACAGCAGCTGTTTCAACATTTGTCTCAAGCAGAAATCCTGAGCCTTAGTG 1707
 Qy 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 Db 1708 TTTGTGTGAAGTACATGCAAGTCTTTTGGCCAAAGTTCACGGTTTATGTGAAGATGTG 1767
 Qy 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 Db 1768 AATGAAACACCTCAATTTTCCCAACAGTATTTCCAAAGCGAAATCACTGATGATGATGCT 1827
 Qy 1828 ATAGGCACTAAAGTGGCAATGATGATGCAAGATCCAGAAAGTCTGACATAGCTAT 1887
 Db 601 SerLeuArgGlyAspThrArgGlyTyrPheLysIleAspHisValThrGlyGluIlePhe 620
 Db 1888 TCACGTGAGGAGACACAAAGAGTGGCTTAAATTTGACCACTGATGCTGTGAGATCTTT 1947
 Qy 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaThr 640
 Db 1948 AGGTGCTCATTTGGACAGAGAGGCGGAGATCCATATGGGTACAAAGTGGTGGCCACA 2007
 Qy 641 GluValGlyLysSerLeuSerValSerGluPheHisLeuIleLeuMetAspVal 660

Db 2008 GAAGTACGGGGGCTTCTTAAAGCTGTGTGACAGTGTCCAGCTGATCCTTATGATGTG 2067
 QY AsnAspAsnProProArgLeuAlaAlaAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 2068 AATGACAAACCTCCCGAGGCTAGCCAAAGACTACACGGGCTGTCTTCTTCCATCCCTC 2127
 QY SerAlaProGlySerLeuIlePheGluAlaThrAspAspGlnHisLeuPheArgGly 700
 Db 2128 AGGACACCTGGAAGTCTCATTTTCGAGCTACTGATGATGATGACACTTATTTCGGGGT 2187
 QY ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerLys 720
 Db 2188 CCCATTATTAATTTTCCTTCGCGAGTGAAGCTTACAAAACAGCTGGGAAGTTTCCAAA 2247
 QY IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaTyr 740
 Db 2248 ATCAATGATCTACTATGCCGACCTGTCTTACACAGCACACACACTTTGAGGAGGGCTAT 2307
 QY ValValLeuIleArgIleAsnAspGlyLysArgProProLeuGluGlyIleValSerLeu 760
 Db 2308 GTGCTCTTGATCCGCATCAATGATGGGGGTGCGCCACCTTGGAAGGCAATTGTTCTTGA 2367
 QY ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
 Db 2368 CCAATTACATTCGACGTTGTGTGAAGAGAGTTGTTTCGGCCACAGGCTCACCAACT 2427
 QY GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2428 GGGATACCCACTGTGGCGATGCGAGTTGGTACTGTGACACCCCTTCTGGGATTTGAT 2487
 QY IleIleLeuAlaValAlaPheIleArgIleLysLysAspLysGlyLysAspAsnValGlu 820
 Db 2488 ATATTTTATGACAGTTGGTTTATCCGCATTAAGAGATTAAGCAAAAGATTAATGTTGAA 2547
 QY SerAlaGlnAlaSerGluValLysProLeuArgSer 832
 Db 2548 AGTCTCAAGCATCTGAAGTCAAACTCTGAGAAGC 2583

RESULT 2

US-09-922-217-1076
 ; Sequence 1076, Application US/09922217
 ; Patent No. US2002076414A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Secrist, Heather
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Meagher, Madeleine Joy
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Wang, Tonglong
 ; APPLICANT: Jiang, Yugu
 ; APPLICANT: Smith, Carole Lynn
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Clapper, Jonathan D.
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
 ; FILE REFERENCE: 210121.471C13
 ; CURRENT APPLICATION NUMBER: US/09/922.217
 ; NUMBER OF SEQ ID NOS: 1124
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1076
 ; LENGTH: 3345
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-922-217-1076

Alignment Scores:

Pred. No.: 0 Length: 3345
 Score: 4321.00 Matches: 832
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0
 US-10-025-380-1081 (1-832) x US-09-922-217-1076 (1-3345)

QY 1 MetIleLeuGlnAlaHisLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
 Db 88 ATGATCTTCAAGGCCCATCTTCACTCCCTGTGTCTTATGCTTATTTGGCAACTGGA 147
 QY 21 TyrGlyGluGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGlu 40
 Db 148 TATGSCCAAGGGGAAAGTTTATGTGACCCCTGAAAACCAATGACATTTCTATTATGAA 207
 QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheValAlaAsnProProAlaValThrPhe 60
 Db 208 GGCCAAGAACCGATCAATTAATTCAGATTAAAGCCCAATCTCTGCTGACTTTT 267
 QY 61 GluLeuThrGlyLysIleThrAspAsnIlePheValIleGluArgGluGlyLeuLeuTyrTyr 80
 Db 268 GAACTAACCTGGGGAGACAGCAACATATTGTTGATGAACGGGAGGACTTGTGATTAC 327
 QY 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 Db 328 AACAGAGCCTTGACAGGAGAAACAAATCTACTACAAATCTCCAGGTTGACGCCCTGAC 387
 QY 101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValLysAspIleAsn 120
 Db 388 GCTAATGAAATTAATGAGAGGCTCCAGTCCCTTAATCAATGAATGAAGACATAC 447
 QY 121 AspAsnArgProThrPheLeuGlnSerLysTyrGluGlySerValArgGlnAsnSerArg 140
 Db 448 GACAAATGACCCACGTTTCTCCAGTCAAGTACGAAGGCTCAGTAAGGCAACTCTCCG 507
 QY 141 ProGlyLysProPheLeuTyrValAlaAlaThrAspLeuAspAspProAlaThrProAsn 160
 Db 508 CCAAGAAACCCCTTCTGTATGCAATGCAACAGACACTGATGATCCGACCTCCCAAT 567
 QY 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
 Db 568 GGCACCTTATTAACCAATGTGATCCAGCTTCCATATCAACATGATGATCTT 627
 QY 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
 Db 628 CAGATCAACAAACAAACGGAGCATCTCTTACCAGAGGAGACTCAGGAATTGAAT 687
 QY 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysPheMetGlyGlyGlnSer 220
 Db 688 CCTGCTAAAGATCTCTTATATCTGTGATCTCAGTAAGACATGGAGGCCAGAGT 747
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrLys 240
 Db 748 GAGAAATTTCTTCAAGTATCCACATCTGTGAAATCATATGACACAGAAATTTTGGAAA 807
 QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
 Db 808 GACACCAAAACCTGTGAGATGTGAGAAACTCACTGATCCCTCACCCCATCAAAATCACT 867
 QY 261 GlnValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGluLysLeuPro 280
 Db 868 CAGTCCCGGTGAAATGATCCCGTGACAAATTTCTTGTGTGACAAAGAAAGCTGCCA 927
 QY 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrValThrGlnProLeuAspArg 300
 Db 928 AGATTCCTCATTTTCAATTAACAGAGAGAAATTTACGTACCTAGCCCTTGACCA 987
 QY 301 GluGluLysAspAlaTyrValPheTyrAlaValAlaAlaAspGluTyrGlyLysProLeu 320
 Db 988 GAAGAAAGAGATGATATGTTTTTATGCAAGTTGCAAGAGATGAGAGAAACCACTT 1047
 QY 321 SerTyrProLeuGluIleHisValLysValLysAspIleAsnAspAsnProThrCys 340
 Db 1048 TCATATCCCGCTGAAATTCATGTAAGTTAAAGTAAATTAAGATATTCACCTCATGCT 1107

QY 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 DB 1108 CCGTCACCAAGTACCGGATTTGAGTCCAGAGAAATACGATCGGTAACAGATCGG 1167
 QY 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
 DB 1168 ACCCTACTGCACATGACAGGAGTGAAGAAATATCTCCAAACAGTTTCTTAATCACTACAG 1227
 QY 381 IleValGlnGlnThrProGlyLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 DB 1228 ATGTGTGAGCAAACTCCCAAACTTCCATGATGAGTCTTCTTAATCCAAACCTATGCT 1287
 QY 401 GluMetLeuGlnLeuAlaValGlnSerLeuLysGlnAspThrProGlnTyrAsnLeu 420
 DB 1288 GGAATGTACAGTTAGCTTAACAGTCTTGAAGAGCAAGTACTCTCTGATCAACTTA 1347
 QY 421 ThrIleGluValSerAspLysAspPheLysThrLeuGlyPheValGlnIleAsnValIle 440
 DB 1348 ACGATGAGGTGTGTGCAAAAGATTTCAAGACCTTTGTTTGTGCAAAATCAACGTATAT 1407
 QY 441 AspIleAsnAspGlnIleProIlePheGlnLysSerAspTyrGlyAsnLeuThrLeuAla 460
 DB 1408 GATATCATATGATGATGATCCCATCTTGAAGAAATCAATATGAGAACTGACTCTGCT 1467
 QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 DB 1468 GAAAGCAACAAACATTTGGGTCCACATCTTAACCTCAAGGCCACTGATGCTGATGAGCCA 1527
 QY 481 PheThrGlySerSerLysIleLeuTyrHisIleIleLysGlyAspSerGluGlyArgLeu 500
 DB 1528 TTTACTGGAGTCTTAATAATCTGTATCATCATTAAGGAGACAGTGAAGGAGCGCTG 1587
 QY 501 GlyValAspThrAspProHisThrAspThrGlyTyrValIleIleLysLysProLeuAsp 520
 DB 1588 GGGGTTCACACAGATCCCAACCAACCAACCGAATATGATTAATAAAGCCCTTGAT 1647
 QY 521 PheGluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 DB 1648 TTTGAAACAGACGCTGTTTCCACATCTGTGTTCAACAGAAATCTTGAGCCCTGATG 1707
 QY 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 DB 1708 TTTGCTGTGAAGTACATGCAAGTTCTTTTCCAACTTCAAGCTTATTTGACAGATGTG 1767
 QY 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 DB 1768 AATGACACACCTCAATTTTCCCAACAGATATCCAAACGAAAGTCACTGAGATGTAGCT 1827
 QY 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTyr 600
 DB 1828 ATAGGCACTAAAGTGGGCAATGTATGCTCCAAAGATCCAAAGGCTTGACATTAAGCTAT 1887
 QY 601 SerLeuArgGlyAspThrArgGlyTyrPheLysIleAspHisValIleGlyGlnIlePhe 620
 DB 1888 TCACTGAGGGGACACACAGAGGTTGGCTTAATTAATGACCACTGAGTGAAGATCTTT 1947
 QY 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaIleThr 640
 DB 1948 AGTGTGCTCCATTTGACAGAGAGCGGAAGTCCATATCGGTACAGAGTGTGCACCA 2007
 QY 641 GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 DB 2008 GAAATGAGGGGGGTCTTCTTAAGCTCTGTGTGCAAGTTCCACTGATCCCTTAATGAGATGTG 2067
 QY 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 DB 2068 AATGACAACTCCCAAGGCTAGCAAGGACTACACGGGCTGTTCTTGCCATCCCTC 2127
 QY 681 SerLysProGlySerLeuIlePheGluAlaThrAspAspArgGlnHisLeuPheArgGly 700
 DB 2128 AGTGCACCTGGAATCTCATTTTGCAGGCTACTGATGATGATGATGACTTAATTTTGGGGT 2187
 QY 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerLys 720

DB 2188 CCCCATTTTACATTTTCCCTCGCAGTGAACCTTCAAAAGACAGTGGAAATTTCCAAA 2247
 QY 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisIleThrAspPheGluGluArgAlaTyr 740
 DB 2248 ATCAATGTATCACTCAAGCCGACGTGTACAGGACACAGACTTGAAGAGGGGGAT 2307
 QY 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGlnGluIleValSerLeu 760
 DB 2308 GTGCTTTATCCGATCAATGATGAGGGGTCCGCCACCTTGGAAAGCATTTGTTCTTAA 2367
 QY 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
 DB 2368 CCAATTACATTTGCACTGTGTGTGGAAGAAAGTTGTTCCGGCCAGACAGTCAACAGACT 2427
 QY 781 GlyIleProThrValGlyMetAlaValIleLeuLeuThrThrLeuLeuValIleGly 800
 DB 2428 GGGATACCACTGTGGCATGAGGAGTGTGATACCTCAACCCCTTGTGTGATGAT 2487
 QY 801 IleIleLeuAlaValAlaPheIleArgIleLysLysAspLysGlyLysAspAsnValGlu 820
 DB 2488 ATATTTTACAGCTGTGTGTTATCCCATTAAGAAAGATTAAGGCAAAAGTATATGTGA 2547
 QY 821 SerAlaGlnAlaSerGluValLysProLeuArgSer 832
 DB 2548 AGTGTCAAGCATCTGAAGTCAAACTCTGAGAGAC 2583

RESULT 3
 US-09-962-436-302
 ; Sequence 302, Application US/09962436
 ; Patent No. US20020081301A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Soppel, Daniel
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
 ; FILE REFERENCE: 689290-75
 ; CURRENT APPLICATION NUMBER: US/09/962,436
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,082
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/234,924
 ; NUMBER OF SEQ ID NOS: 568
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 302
 ; LENGTH: 3345
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-962-436-302

Alignment Scores:
 Pred. No.: 0 Length: 3345
 Score: 4321.00 Matches: 832
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-10-025-380-1081 (1-832) x US-09-962-436-302 (1-3345)
 QY 1 MetIleLeuGlnAlaHisLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
 DB 88 ATATATCTTACAGCCCATCTTCACTCCCTGTGTCTTATGCTTATTTAGGCACTGGA 147
 QY 21 TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGlu 40
 DB 148 TATGGCCCAAGAGGGGAAGTTATGTGACCCCTGAAACCATGATCAATTTCTATTATGAA 207
 QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheValAlaAsnProProAlaValThrPhe 60
 DB 208 GGCACAAAGACCGAGTCAATATATATTCAGTTTAAGGCCATATCTCTGCTGACCTTTT 267
 QY 61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluArgGluGlyLeuLeuTyrTyr 80

```

Db      268 GAACATACTGGGGAGACACACATATTTGTGATAGAACGGGGAGGACTTCTATATAC 327
Qy      81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGluValAlaAlaLeuAsp 100
Db      328 AACGAGCCTTTGGACAGGAAACAGATCTTACTACAAATCTCCAGGTGGACCCCTGGAC 387
Qy      101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValIleAspIleAsn 120
Db      388 GCTATAGGAATTAATAGGAGGAGGTCCAGTCCCTATACACATAGAAAGGAGGACATCAAC 447
Qy      121 AspAsnArgProThrPheLeuGlnSerLysGlyGluGlySerValIleArgIleAsnSerArg 140
Db      448 GACATATGACCCACGTTCTCCAGTCAAGTACGAAAGCTCAGTAAAGGAGAACTCTGCG 507
Qy      141 ProGlyLysProPheLeuArgValAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
Db      508 CAAAGAAAGCCCTCTGTGATGTCAATGCAAGACCAAGCTGGATGATCCGGCCACTCCCAAT 567
Qy      161 GlyGlnLeuArgTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
Db      568 GGGCAGCTTTATACAGATTTGTATCATCAGCTTCCATGATCAACAATGTCTATCTT 627
Qy      181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
Db      628 CAGATCAACAACAAACGAGGAGCATCTCTTACCCGAGAGGGATCTCAGGAATTGAAT 687
Qy      201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnIleSer 220
Db      688 CCTGTCAAGAACTCTCTCTAATCTGTATCTGTATCTCAAGTGAAGGACATGGAGGCAAGT 747
Qy      221 GluAsnSerPheSerAspThrThrSerValAspIleIleValIleThrGluAsnIleTyrLys 240
Db      748 GAGATTCCTTCACTGATACCAATCTGTGTGATATCTAGTGAAGAGAAATTTGGAAA 807
Qy      241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
Db      808 GCACCAAAACCTGTGAGATGTGTGAAAACTCAACATGATCTCCATCAACCAAAATCACT 867
Qy      261 GluValAlaArgThrAsnAspProGlyValGlnTyrSerLeuValAspLysGluLysLeuPro 280
Db      868 CAGGTGGGTGAGAAATGATCCCGGTGCAAAATATCTTACTGTGCAAAAGAAACCTGCCA 927
Qy      281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrValIleThrGlnProLeuAspArg 300
Db      928 AGATTCACATTTTCAATTGACACGAGAGAGATTTTACGTGACTCAGCCCTTGGACGA 987
Qy      301 GluGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu 320
Db      988 GAAAGAAAGATGATATGTTTATTTTATGCAAGTTGCAAGATGATGAGTAAACCACTT 1047
Qy      321 SerTyrProLeuGluIleHisValLysValLysAspIleAsnAspAsnProProThrCys 340
Db      1048 TCATATCTCGCTGAAATTCATGTAAAGTTAAAGATTTAATGATTAATCCACCTACATGCT 1107
Qy      341 ProSerProValThrValPheGluValGlnGluAsnGluLysLeuGlyAsnSerIleGly 360
Db      1108 CCGTCAACCAATACCGATTTTGAAGTCCAGAGAAATACAGACTGGGTAAACAGTATGGG 1167
Qy      361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
Db      1168 ACCCTACTGCACATGACAGGATGAGAGAAATACCTCCAAACAGTTTCTTAACTACAGG 1227
Qy      381 IleValGluGlnIleThrProLysLeuProMetAspGlyLeuPheLeuIleGlnIleThrVal 400
Db      1228 ATGTGTGAGCAAACTCCCAAACTTCCATGATGAGATCTTCCTAATCCAAACCTTAAGCT 1287
Qy      401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyrAsnLeu 420
Db      1288 GGAATGTTACAGTTAGTAAACAGTCTCTTGAAGAAGACAGATACTCTCTCAGTACAACTTA 1347
Qy      421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
Db      1348 ACGATAGAGGTGTCTGACAAAGATTTCAAGACCCCTTTGTGTGCAAAATCAAGTTATT 1407
Qy      441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAla 460
Db      1408 GATTTCAATATCATGATCCCATCTTTGAAAAATCATATTAAGAAACCTGACTTTTGTCT 1467
Qy      461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
Db      1468 GAAAGACAAACATTTGGGTCCACCATCTTAACCATCCAGGCCACTGATGTGATGACCA 1527
Qy      481 PheThrGlySerSerLysIleLeuTyrHisIleIleLysGlyAspSerGluGlyValLeu 500
Db      1528 TTTACTGGGAGTTCTAAATTTCTGTATCATATATCAATAAAGGAGACAGTGAAGGCGCTG 1587
Qy      501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
Db      1588 GGGGTTCACAGATCCCAATCCCAACACGAGAAATGTCATTAATTAAGACCTCTGTAT 1647
Qy      521 PheGluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
Db      1648 TTTGAAACACAGCTGTTTCCAAATTTGTCTTCAAGACAGAAATCTCGAGCTTATGTG 1707
Qy      541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValIleAspVal 560
Db      1708 TTTGTGTGAAAGTACAAATGCAATGCTTTTGGCAAGTTCAAGCTTAATTTGACAGATGTG 1767
Qy      561 AsnGluAlaProGlnPheSerGlnHisValPheGluAlaLysValSerGluAspValAla 580
Db      1768 AATGAAGCACCTCAATTTTCCAAACAGTATTCAGAGGAAAGTCAATGAGAGATGAGCT 1827
Qy      581 IleGlyThrLysValGlyAsnValIleThrAlaLysAspProGluGlyLeuAspIleSerTyr 600
Db      1828 ATAGGACATTAAGTGGGCAATGTGACTGCCAAGATCCAAAGTCTGACATTAAGCTAT 1887
Qy      601 SerLeuArgGlyAspThrArgGlyTyrPheLysIleAspHisValIleThrGlyLysIlePhe 620
Db      1888 TCACGTGAGGAGACCAAGAGGTTGGCTTAAATATGACCAAGCTGATCTGTGATCTTT 1947
Qy      621 SerValAlaProLeuAspArgGluAlaGlySerProTyrTyrValGlnValValAlaThr 640
Db      1948 AGTGTGCTTCATTTGACAGAGAGCCGGAAGTCCATATCGGTGATCAAGTGTGGCCACA 2007
Qy      641 GluValGlyLysSerSerLeuSerSerValSerGlyPheHisLeuIleLeuMetAspVal 660
Db      2008 GAAGTAGGGGGGCTTCTTAACTCTGTCTCAGAGTCCACCTGATCTTAATGATGTG 2067
Qy      661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
Db      2068 AATGACAAACCTCCAGGCTAGCCAAAGACTACAGGGCTGTGTTCTTCCATCCCTC 2127
Qy      681 SerIleAspProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheAspGly 700
Db      2128 AGTGCACCTGGAGGTCTCATTTTTCAGGCTACAGATGATATAGACATTAATTCGGGGT 2187
Qy      701 ProHisPheThrPheSerLeuGlySerGlySerLeuGluAsnAspTyrGluValSerLys 720
Db      2188 CCCCATTTTACATTTTCCCTCGGCACTGAGAACTTACAAACACATCGGAAATTTCCAA 2247
Qy      721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluValArgAlaTyr 740
Db      2248 ATCAATAGTACTCATCGCCGACATGTCTACACAGGACACACAGACTTTGAGAGAGGGCTAT 2307
Qy      741 ValValLeuIleArgIleAsnAspGlyLysArgProProLeuGlnGluGlyIleValSerLeu 760
Db      2308 GTGCTGTGATTCGATCAATGATGAGGAGGCTGAGCCCTTGAAGGACATTTGTTCTTAA 2367
Qy      761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnIleThr 780
Db      2368 CCAATTAATCTTCGAGTGTGTGTGAGAGAAAGTTTTCGGGCAACAGGTTCACAGACT 2427
Qy      781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
Db      2428 GGGATACCCCACTGTGGGATGAGTGTGATTAATCTGACCAACCTTCTGTGTGATGTGT 2487

```

Qy 801 ILEILELEUAAVALAVALPHEILEARGLIELEYSLSAPLYSGLYLSAASPANVALGLU 820
Db 2488 ATAATTTTAGAGTTGGTTTATCCGCAATAAGCAAGCAATGATATGTTGAA 2547
Qy 821 SerIaGlnAlaSerGluValLysProLeuArgSer 832
Db 2548 AGTCTCAAGCATCTGAAGTCAAACTCTGAGAAAGC 2583

RESULT 4

US-09-833-263-1076
; Sequence 1076, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Mesgher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1076
; LENGTH: 3345
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-263-1076

Alignment Scores:

Pred. No.: 0
Score: 4321.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Gaps: 0

Length: 3345
Matches: 832
Conservative: 0
Mismatch: 0
Indels: 0

US-10-025-380-1081 (1-832) x US-09-833-263-1076 (1-3345)

Qy 1 MetIleuGlnAlaHisIleuHisSerLeuCysLeuLeuMetLeuTYrLeuAlaThrGly 20
Db 88 ATGATCTTCAGGCCCACTTCACTCCCTGCTCTTCTTATGCTTATTTGGCAACTGGA 147
Qy 21 TyGIgInGInGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTYrGlu 40
Db 148 TATGGCCAAAGGGGAGTTAGTGAACCCCTGAAACCCATGACATTTTCTATTTATGAA 207
Qy 41 GlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAsnProProlaValThrPhe 60
Db 208 GGGCAAGAACCGAAGCAATTAATATTCAGTTTAAGGCCAATCCCTCGCTGAGCTTTT 267
Qy 61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluArgGluGlyLeuLeuTYr 80
Db 268 GAATTAAGTGGGAGAGACAGACATATTTGTGATAGAACGGAGGAGCTTCTGATTAAC 327
Qy 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
Db 328 AACAGAGCTTGGACAGGAGAAACAAAGATCTACTCAATCTCCAGGTTCAGCCCTGGAC 387
Qy 101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValLysAspIleAsn 120
Db 388 GCTATAGAAATTAATAGTGAAGGAGTCCAGTCCCTATCACATAGAGTAAAGGACATCAAC 447
Qy 121 AspAsnArgProThrPheLeuGlnSerLysTYrGluGlySerValArgGlnAsnSerArg 140
Db 448 GACATGACACCACTTCTTCAGTCAAGTCAAGTACAGAGGCTCAGTAGGCAAACTCTGCG 507
Qy 141 ProGlyLysProPheLeuTYrValAsnAlaThrAspLeuAspArgProAlaThrProAsn 160
Db 508 CCAGGAAAGCCCTTCTTGTATGTCAATGCCAGACAGACTGAGTATCCGCACTCCCAAT 567

Qy 161 GlyGlnLeuTYrTYrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTYrPhe 180
Db 568 GGGCAGCTTTATACAGATGTCATCCAGCTTCCCATGATCAACATGCTACTTCT 627
Qy 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
Db 628 CAGATCAACAAACAAACCGGAGGCATCTCTTACCCGAGGAGGATCTCAGAAATGTAAT 687
Qy 201 ProAlaLysAsnProSerTYrAsnLeuValIleSerValLysAspMetGlyGlnSer 220
Db 688 CTGCTAAGATCTTCTTATATCTGGTATCTTCATGAGGACATGGAGGCCAGACT 747
Qy 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTYrPlys 240
Db 748 GAGATTTCTTCACTAGTATACCATCTGTGATATCATGATGACAGAAATATTTGAAA 807
Qy 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
Db 808 GCACCAAAACCTGTGAGATGGTGAAACCTCAACTGATCTTCAACCCATCAAAATCCTACT 867
Qy 261 GlnValArgThrAsnAspProGlyValIleGlnTYrSerLeuValAspLysGluLysLeuPro 280
Db 868 CAGGTGGGTGAGATGATCCCGGTGCAATATTTCTTACTGACAAAGAGAGCTGCCA 927
Qy 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTYrValThrGlnProLeuAspArg 300
Db 928 AGATTCACATTTCAATTTGACAGAGAGAGATATTTAGCTGACTCAGCCCTTGACCGCA 987
Qy 301 GluGlnLysAspAlaTYrValPheTYrAlaValAlaLysAspGluTYrGlyLysProLeu 320
Db 988 GAAAGAAAGCATATGTTTATGTCAGTTCAGAAAGATGATGACGAGAAACCACTT 1047
Qy 321 SerTYrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProThrCys 340
Db 1048 TCATATCCGTGGAATTCATGTAAGTAAAGTTAAAGTATATATATCACTACTACTGT 1107
Qy 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
Db 1108 CCGTACACCAATTAACCGATTTTGAAGTCCAGAGAAATGACATCGGGTAACTATTCGGG 1167
Qy 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTYrArg 380
Db 1168 ACCCTTACTGACATGACAGGAGTGAAGAAATACGCAACAGTTTCTTAACTCAGG 1227
Qy 381 IleValGlnGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTYrAla 400
Db 1228 ATGTGTGAGCAAACTCCCAAACTTCCATGATGAGACTCTTCTTAATCAAACTATGCT 1287
Qy 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTYrAsnLeu 420
Db 1288 GGAATGTTACATGATTAACAGTCTTGAAGAGCAATATCTCCTCAGTCAACTTA 1347
Qy 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
Db 1348 ACCATAGAGGTGTGTGCAAAAGATTTCAAAACCTTGTGTTTGTGCAAACTGATAT 1407
Qy 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTYrGlyAsnLeuThrLeuAla 460
Db 1408 GATATCATATATGATGATCCCATCTTGTGAATAATCATATTAAGGAACCTGACTCTTGGCT 1467
Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
Db 1468 GAAGACACAAACATTTGGGTCCACATCTTAAACCATCAGGCCACTGATGATGAGCA 1527
Qy 481 PheThrGlySerSerLysIleLeuTYrHisIleIleLysGlyAspSerGluGlyArgLeu 500
Db 1528 TTTACTGGAGATTCTAAATATCTGTATCATATCATTAAGGAGAACAGTGGAGGACGCTG 1587
Qy 501 GlyValAspThrAspProHisThrAsnThrGlyTYrValIleIleLysLysProLeuAsp 520
Db 1588 GGGGTGTACACAGATCCCATACCAACCGGAAATGTCATTAATTAAGGCTCTTGAT 1647
Qy 521 PheGluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540

Db 1648 TTGAAACAGCAGCTGTTTCCAACTGTTTCAAGAGAGAAATCCGACCTCTAGTG 1707
 Qy PheGlyValIysrYrAsnAlaSerSerPheAlaIysPheThrLeuIleValThrAspVal 560
 Db 1708 TTTGGTGGAGTGCATGATGCAATGTTCTTTTCGCAAGTTCAAGCTTATTGGACAGAGTG 1767
 Qy AsnGluAlaProGlnPheSerGlnIleValPheGlnAlaIysValSerGluAspValAla 580
 Db 1768 AATGAGACCTCAATTTTCCCAACAGTATTCAGAGGAAAGTCAGAGGAGTGTAGCT 1827
 Qy 581 TleGlyThrIysValGlyAsnValThrAlaIysAspProGluGlyLeuAspIleSerTyr 600
 Db 1828 ATAGGCACTAAAGGCGGCAATGTAGTCCCAAGATCCAGAAAGCTCGACATAGCTAT 1887
 Qy SerLeuArgGlyAspThrArgGlyTyrLeuIysIleAspHisValThrGlyGluIlePhe 620
 Db 1888 TCACCTAGGGGAGACACAGAGGTTGGCTTAAATGACCAAGACTGTGAGATCTTT 1947
 Qy SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaThr 640
 Db 1948 AGTGTGCTCCATTTGACACAGAGAGCCGGAAGTCCATTCGGGTACAGAGTGTCGCCACA 2007
 Qy 641 GluValGlyIysSerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 Db 2008 GAAGTAGGGGGGTCTTCTTAAGCTGTGTCAAGATTCACATGCTTATGATGATGTG 2067
 Qy 661 AsnAspAsnProProAspLeuAlaIysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 2068 AATGACAAACCTCCAGAGCTTAGCGAAGACTACAGGGCTGTGTTCTTGTGCAATCCCTC 2127
 Qy 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspArgGlnIleLeuPheArgGly 700
 Db 2128 AGTGCACCTGGAAGTCTCATTTTGAAGGCTACTGATGATGATGACATTAATTCGGGGT 2187
 Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerIys 720
 Db 2188 CCCCATTTTCAATTTTCCCTCGCGCAGTGAAGCTTACAAACGACTGGAAAGTTTCCAA 2247
 Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluAlaThr 740
 Db 2248 ATCAATGATCTACATGCGCCAGCTGTCTACACAGGACACAGCTTGAAGAGGCGCTAT 2307
 Qy 741 ValValLeuIleArgIleAsnAspGlyIysArgProProLeuGluGlyIleValSerLeu 760
 Db 2308 GTGTCTTGAATCCGACATGATGATGAGGGGTCCGCCACCTTGAAGGACATTTGTTCTTA 2367
 Qy 761 ProValThrPheCysSerCysValGlyGlySerCysAspPheArgProAlaGlyHisGlnThr 780
 Db 2368 CCAATTACATTCGACAGTGTGTGAGAGAAAGTTGTTCCGGCCAGCAGGTCACCAACT 2427
 Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2428 GGAATATCCCATGCGGCGCATGCGGTGATATGATGATGACACACCTTCTGGTATTTGGT 2487
 Qy 801 IleIleLeuAlaValPheIleArgIleIysIysAspIysGlyIysAspAsnValGlu 820
 Db 2488 ATAATTTTACAGTGTGTGTTATCCGATTAAGAGATTAAGGACAAAGATATATGTTGAA 2547
 Qy 821 SerAlaGlnAlaSerGluValIysProLeuArgSer 832
 Db 2548 AGTGCTCAAGCAATCTGAAGTCAAACTCTGAAGAGC 2583
 RESULT 5
 US-09-880-107-3288
 ; Sequence 3288, Application US/09880107
 ; Patent No. US20020142981A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Horne, Darci T.
 ; APPLICANT: Vockley, Joseph G.
 ; APPLICANT: Scherf, Uwe
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

; FILE REFERENCE: 44921-5028-WO
 ; CURRENT APPLICATION NUMBER: US/09/880,107
 ; CURRENT FILING DATE: 2001-06-14
 ; PRIOR APPLICATION NUMBER: US 60/211,379
 ; PRIOR FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: US 60/237,054
 ; PRIOR FILING DATE: 2000-10-02
 ; NUMBER OF SEQ ID NOS: 3950
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3288
 ; LENGTH: 3345
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURES:
 ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U07969
 US-09-880-107-3288
 Alignment Scores:
 Pred. No.: 0 Length: 3345
 Score: 4321.00 Matches: 832
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0
 US-10-025-380-1081 (1-832) x US-09-880-107-3288 (1-3345)
 Qy 1 MetIleleuGlnAlaHisIleLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
 Db 88 ATATATCTTACAGGCGCCATCTTCACTCCCTGTGCTTTATATGTTATTTGGCAACCTGGA 147
 Qy 21 TyrGlyGlnGluGlyIysPheSerGlyProLeuIysProMetThrPheSerIleTyrGlu 40
 Db 148 TATGGCCAGAGGGGAGAGTTTATGAGACCTTGAAACCCATGACATTTTCTATTTATGAA 207
 Qy 41 GlyGlnGluProSerGlnIleIlePheGlnPheIysAlaAsnProProAlaValThrPhe 60
 Db 208 GGCACAAAGACCGAGTCAATATATATCCAGTTTAAGGCCAATCTCTGCTGTGACTTTT 267
 Qy 61 GluLeuThrGlyIysThrAspAsnIlePheValIleGluArgGluGlyLeuTyrTyr 80
 Db 268 GAACATACTGGGAGAGACAGCAACATATTTGTATGATGAACGGAGGAGCTTCTGTATTAC 327
 Qy 81 AsnArgAlaLeuAsnAspArgIuThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 Db 328 AACAGACCTTGGACAGGGAGAAACAGATCTTACATCTCAATCTCAAGGTGCGACTGGAC 387
 Qy 101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValIysAspIleAsn 120
 Db 388 GCTAATGGAATTTATGATGAGGGGTCCAGTCCCTATCACCATAGAGAGGACATCAAC 447
 Qy 121 AspAsnArgProThrPheLeuGlnSerIysTyrGlyIysValIysArgGlnAsnSerArg 140
 Db 448 GACAAATGCAACCCAGTTCCTCCAGTCAAGTAAGAGGCTCAATAAGGCGAGAACTCTGCG 507
 Qy 141 ProGlyIysProPheLeuTyrValIleAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
 Db 508 CAGAGAAAGCCCTTCTTGTATGATGATGACACAGACCTGATATATCCGGCCACTCCCAAT 567
 Qy 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
 Db 568 GGCACCTTATATACAGATGTCATCCAGCTTCCCATATGACAGATGTCATGATCTTT 627
 Qy 181 GlnIleAsnAsnIysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
 Db 628 CAGATCAACAAACAAACGGAGGCATCTCTTACCCGAGAGGATCTCAGGAATTTGAAT 687
 Qy 201 ProAlaIysAsnProSerTyrAsnLeuValIleSerValIysAspMetGlyGlyIys 220
 Db 688 CTGCTAAGATCTCTTCTTATATCTGAGTATTCAGTGAAGACATGGAGGCCAGAGT 747
 Qy 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrPlys 240

Db 748 GAGAAATTCCTTCAGTGAATACCAACATCTGTGATATCATAGTACAGAGAAATATTTGGAAA 807
 Qy 241 AAlaProlyserProValGluMetValGluAsnSerThrAspProHisProIleValIleThr 260
 Db 808 GCACCAAAACCTGGAGATGTGTGAAAACCACTGATCTCTCAACCCATCAAAATACT 867
 Qy 261 GlnValArgThrAsnAspProGlyValGlnTyrSerLeuValAspLysGluLysLeuPro 280
 Db 868 CAGGTGGGGTGAATGATCCCGGTGCACAAATTTCTTGAATTGACAAAGAACTCCCA 927
 Qy 281 ArgPheProPheSerIleAspGlnGluValAspIleTyrValThrGlnProLeuAspArg 300
 Db 928 AGATTCCCATTTTCATATGACCAAGAGATATTTACGACTCAGCCCTTGAGACGA 987
 Qy 301 GlnGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu 320
 Db 988 GAAAGAAAAGATGATATGTTTTTATGCAAGTTCAAAGATGAGTACGAGAAACCACTT 1047
 Qy 321 SerTyrProLeuGluIleHisValIleValLysAspIleAsnAspAsnProProThrCys 340
 Db 1048 TCATATCCGCTGGAAATTCATGTAAGTTAAAGTATATATGATATATCCACTACATGT 1107
 Qy 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 Db 1108 CCGTACCAAGTAAACCGATTTGAGGTCCAGAGATGAACGATGGGTAAACAGATATCGGG 1167
 Qy 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
 Db 1168 ACCCTTACTGCACATGACAGGATGAAAGAAATACTGCAACATTTCTTAACTACAGG 1227
 Qy 381 IleValGlnGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 Db 1228 ATTGTGAGCAAACTCCCAACTTCCCATGATGATGATCTCTTCAATCCAACTATGCT 1287
 Qy 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyrAsnLeu 420
 Db 1288 GGAATGTATACAGTATGCTAAACAGTCTCTTGAAGAACAAAGATACCTCCATCACTAACTTA 1347
 Qy 421 ThrIleGluValSerAspLysAspPheLysPheLysCysPheValGlnIleAsnValIle 440
 Db 1348 ACGATAGAGGTGCTGACAAAGATTTCAAGACCTTGTGTTTGTGCAATCAAGTATAT 1407
 Qy 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAla 460
 Db 1408 GATATCAATGATTCAGATCCCATCTTTGAAAAATCAGTTATGAAACCTCACTTGGCT 1467
 Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db 1468 GAAAGACAAACATTTGGGTCCACCATCTTAACCATCCAGGCCACTGATGCTGATGAGCA 1527
 Qy 481 PheThrGlySerSerLysIleLeuTyrHisIleIleLysGlyAspSerGluGlyArgLeu 500
 Db 1528 TTTTACTGGAGATTTCTAAATCTGTATCATATCAATAAGGAGACAGAGAGGAGCGCTG 1587
 Qy 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
 Db 1588 GGGGTGACACAGATCCCATACCAACACCGGATHTGCAATATTAATAAACCTCTTGAT 1647
 Qy 521 PheGluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 Db 1648 TTTGAAACAGACAGCTGTTTCAACATTTGTGTTCAAGAGAGAAATCCCTGAGCCCTAGTG 1707
 Qy 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 Db 1708 TTTGTGTGAGATCAATGCAAGTCTTTTCCCAAGTTCACAGCTTATTTGTACAGAGTGTG 1767
 Qy 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 Db 1768 AATGAGCACTCATTTTCCCAACACGATATTCAGAGGAAAGTCAGAGAGATGTACT 1827
 Qy 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTyr 600
 Db 1828 ATAGGCACTAAAGTGGGCAATGTGACTGCCAAGGATCCAGAAAGTCTGACATATAGCTAT 1887

Qy 601 SerLeuArgGlyAspThrArgGlyTyrPheLysIleAspHisValThrGlyGluIlePhe 620
 Db 1888 TCATGAGGGGAGACACMAAGGTGTGCTTAAATGATACCACTGATGATGATCTTT 1947
 Qy 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaIleThr 640
 Db 1948 AGTGTGGCTCATTTGACAGAGAACCGGAAAGTCCATATGCGGTATACAGTGTGGCCACA 2007
 Qy 641 GluValGlyLysSerLeuSerSerValSerGluPheHisIleuIleuMetAspVal 660
 Db 2008 GAAGTGGGGGCTCTTCTTAAGCTCTGTCTCAGAGTTCACCTGATCTTATGAGATGTG 2067
 Qy 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 2068 AATGACAACTCTCCAGGCTAGCCAGCAAGACTACACGGGCTGTCTTCTGCAATCCCTC 2127
 Qy 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheArgGly 700
 Db 2128 AGTGCACCTGGAAGTCTCATTTTTCGAGGCTACGATGATGATCAGACATTAATTCGGGGT 2187
 Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAspAspThrProGluValSerLys 720
 Db 2188 CCCCATTTTACATTTTCCCTCGGACGTGAAAGCTTCAAAAACGACTGGGAAATTTCCAAA 2247
 Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluAlaTyr 740
 Db 2248 ATCAATGTACTCTATGCCCCGACCTGTCTACCGGACACACAGACTTTAGAGAGAGGCTAT 2307
 Qy 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGlnGlyIleValSerLeu 760
 Db 2308 GTCGTCTTATCCGATCATATGATGGGGGTGGCCACCTTGAAAGGATTTGTTCTTTA 2367
 Qy 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
 Db 2368 CCAATTACATTCGACGTTGTGTGGAAGATGTTGTTCCGGCAGAGGCTCACAGACT 2427
 Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2428 GGGATACCCACTGTGGGCAATGGCAATGGATCTGACACACCTTGGATGATTCGT 2487
 Qy 801 IleIleLeuAlaValAlaPheIleArgIleLysLysAspLysGlyLysAspAsnValGlu 820
 Db 2488 ATAAATTTAGCAGTGTGTTATCCCATTAAGAAAGATTAAGCAAAAGATATATGTTGAA 2547
 Qy 821 SerAlaGlnAlaSerGluValLysProLeuArgSer 832
 Db 2548 AGTGTCTCAACATCTGAATCAAAACCTTGAGAGAC 2583

RESULT 6
 US-09-962-832-118
 Sequence 118, Application US/09962832
 Patent No. US20020110821A1
 GENERAL INFORMATION:
 APPLICANT: Ebner, Reinhard
 TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signat
 TITLE OR INVENTION: Sets
 FILE REFERENCE: 689290-74
 CURRENT APPLICATION NUMBER: US/09/962,832
 CURRENT FILING DATE: 2001-09-25
 PRIOR APPLICATION NUMBER: US/60/235,077
 PRIOR FILING DATE: 2000-09-25
 PRIOR APPLICATION NUMBER: US/60/235,280
 PRIOR FILING DATE: 2000-09-25
 NUMBER OF SEQ ID NOS: 259
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 118
 LENGTH: 3654
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-962-832-118
 Alignment Scores:

Db 2133 AGTGCACCTGGAGTCTCATTTTCGAGCTACTGATGATGACGACTTATTTCCGGGT 2192
Qy 701 ProhibitinPhosphoserineGlySerGlySerleuGlnAsnAspTyrGluValSerLys 720
Db 2193 CCCATTTCATTTCCCTCGCGAGGAGGCTTACAAAACGACTGGGAAGTTCCAAA 2252
Qy 721 ILeAsnGlyThrHisAlaArgLeuSerThrArgHisAspPheGluGluArgAlaTyr 740
Db 2253 ATCAATGCTACTGCGCCGACTGCTTACACGACACAGAGTTTGAGAGAGGAGGTAT 2312
Qy 741 ValValLeuLeuLeuArgIleAsnAspGlyGlyArgProProleuGluGlyIleValSerLeu 760
Db 2313 GTCCTCTTGATCCGCATCAATGAGGGGGTGGCACTTGGAAAGCATTTGTTCTTTA 2372
Qy 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
Db 2373 CCAAGTTCATTTCTGCGACTTGTGTGGAGGAAGTTGTTCCGCGACGACGTCACCAACT 2432
Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuValIleGly 800
Db 2433 GGGATACCCACTGGGCGATGGCGATTGGTATCTGCTGACCACTTCTGTGATTTGCT 2492
Qy 801 IleIleLeuAlaValIlePheIleArgIleLysLysAspLysGlyLysAspAsnValGlu 820
Db 2493 ATAAATTTAGCAGTGTGTTTATCCGCATAAAGAGATTAAGCAAAAGATTAATGTTGA 2552
Qy 821 SerAlaGlnAlaSerGluValLysProLeuArgSer 832
Db 2553 AGTCTCAGACATCTGAAGTCAAACCTCTGAGAAGC 2588

RESULT 7

US-09-880-107-3807
; Sequence 3807, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-MO
; CURRENT APPLICATION NUMBER: US/09/880.107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3807
; LENGTH: 3654
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X83228
US-09-880-107-3807

Alignment Scores:

Pred. No.: 0
Score: 4303.00
Percent Similarity: 99.76%
Best Local Similarity: 99.52%
Query Match: 99.58%
DB: 10

Length: 3654
Matches: 828
Conservative: 2
Mismatch: 2
Indels: 0
Gaps: 0

US-10-025-380-1081 (1-832) x US-09-880-107-3807 (1-3654)

Qy 1 MetIleuGlnAlaHisLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
Db 93 ATGATTAATTCAGGCGCCATCTTCACTCCCTGTCTTTATGCTTTATTTGGCACTGA 152
Qy 21 TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGlu 40

Db 153 TATGGCAAGAGGGGAAGTTTACTGACCCCTGAACCCATGACATTTCTATTATGAA 212
Qy 41 GlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAsnProProAlaValThrPhe 60
Db 213 GGCAGAAAGACCGAGTCAATTAATTCAGGTTTAAAGGCCAATCTCTGCTGACTTT 272
Qy 61 GluLeuThrGlyGluThrAspAsnIlePheValIleGlyLysGlyLysLeuLysTyr 80
Db 273 GAACTTAAGGAGGAGACAGACAAATATTTGTGATGAAAGGAGGAGGACTTGTATAC 332
Qy 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
Db 333 AACAGAGCCTTGACAGAGGAAACAGATCTACTCAATCTCCAGGTTGACGCCCTGAC 392
Qy 101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValLysAspIleAsn 120
Db 393 GCTAAATGGAATTAATGAGAGGCTCAGTCCCTATACCATTAAGTAAAGGAGACATCAC 452
Qy 121 AspAsnArgProThrPheLeuGlnSerLysTyrGlySerValArgIleAsnSerArg 140
Db 453 GACAAATGACCCACCTTCTCCAGTCAAGTAAGAGGCTCAATAGGAGAACTCTGCG 512
Qy 141 ProGlyLysProPheLeuTyrValAlaAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
Db 513 CCAAGAAAGCCCTTCTTGATATGCAATGCCACAGACTGGATGATCCGCGCACTCCCAAT 572
Qy 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnValMetTyrPhe 180
Db 573 GGCACCTTAATTAACAGATGTCATCCAGTCCAGTCCATGATGCAATGTCATGACTTT 632
Qy 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
Db 633 CAGATCAACAAACAAACGGAGCATCTCTTACCCGAGAGGATCTCAGAAATGAA 692
Qy 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnSer 220
Db 693 CTGCTAAGAAATCCTTCTTAATCTGATCTGATCTCAGTCAAGACATGGAGGCGCAGAT 752
Qy 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrLys 240
Db 753 GAAATATCTCTGATGATACATCATCTGTGATATCATGATGACAGAAATATTTGAAA 812
Qy 241 AlaProLysProValGluMetValGluAsnSerThrAspProIleLysIleThr 260
Db 813 GCACCAAAACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 872
Qy 261 GlnValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGlyLysLeuPro 280
Db 873 CAGGTGGGTGGAATGATCCCGGTGCACAATATTCCTAGTTGACAAAGAGAGCTGCCA 932
Qy 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrValThrGlnProLeuAspArg 300
Db 933 AGATTCATTTCAATGATGACGAGAGGAGATTAATTAATTAATTAATTAATTAATTAAT 992
Qy 301 GluGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu 320
Db 993 GAAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1052
Qy 321 SerTyrProLeuGluIleHisValLysValLysAspIleAsnAspAsnProProThrCys 340
Db 1053 TCATATCCGCTGGAATATCTGTAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 1112
Qy 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyLysSerIleGly 360
Db 1113 CCGTCAACAGTAAACGATTTAGGTCAGAGAAATGAAGATGAGATGAGATGAGATGAGAT 1172
Qy 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
Db 1173 ACCCTTAATCAATGACAGGAGATGAAGAAATATCTCCAAAGTTTCTTAATCAAGG 1232
Qy 381 IleValGluGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400

Db 1233 ATTGTGACGAACTCCCAAACTTCCCATGATGAGTCTTCTTAATCCAAACCTATGCT 1292
 Qy 401 GlyMetLeuGlnLeuAlaValSerLeuLysGlnAspThrProGlnIYAsnLeu 420
 Db 1293 GGATGTTCAGTTCAGTAAACACTCCTGTGAAGAGCAAGTACTCCTCATGACACTTA 1352
 Qy 421 ThrIleGlnValSerAspLysAspPheLysThrLeuGlyPheValGlnIleAsnValIle 440
 Db 1353 ACGATACAGGTCGTGCAAAAGATTTCAGACCTTTGTTTGGCAATCAAGCTTATT 1412
 Qy 441 AsnIleAsnAspGlnIleProIlePheGlnLysSerAspThrGlyAsnLeuThrLeuAla 460
 Db 1413 GAATTCATGATTCAGACCCCATCTTGAATAATCAATTAAGAAACCTGACCTTGTCT 1472
 Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db 1473 GAAGACACAAACATTTGGGTCCACCATCTTAACCTCAAGGCACTGATGCTGAAGACCA 1532
 Qy 481 PheThrGlySerSerLysIleLeuThrIleIleIleLysGlyAspSerGlnGlyArgLeu 500
 Db 1533 TTTACTGGAGATCTTAATAATTTCTGTATCATATCAATAAGGAGACAGTGAGGACGCTG 1592
 Qy 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleIleLysLysProLeuAsp 520
 Db 1593 GGGGTTCACACAGATCCCATACCAACACCGGATATGTCATTAATTAAGCCCTTGAT 1652
 Qy 521 PheGlnThrAlaIleValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 Db 1653 TTTGAAACACACAGCTGTTCCACATTTGTCTTCAAGACAAATACTTGAGCTTATGTC 1712
 Qy 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValIleAspVal 560
 Db 1713 TTTGGTGTGAAGTACATGCAAGTTCTTTGCCAAGTTACGCTTATGTGACAGATGTC 1772
 Qy 561 AsnGlnAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 Db 1773 AATGAAGCACCTCAATTTCCCAACAGTATTCACACGAAAGTACAGTGGATGTGATGCT 1832
 Qy 581 IleGlyThrIleValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTyr 600
 Db 1833 ATAGGCACTAAAGTGGCAATGACATGACGCCAAGATCCAAAGGTCCTGGACATAGCTAT 1892
 Qy 601 SerLeuArgLysAspThrArgLysTyrPheLysIleAspHisValThrGlyIleLysPhe 620
 Db 1893 TCACTGAGGGAGACACAAAGGTTGGCTTAATTAATGACACGAGTGGTGGATGTTT 1952
 Qy 621 SerValAlaProLeuAspArgLysAlaGlySerProTyrArgValGlnValAlaIleThr 640
 Db 1953 AGTGTGGCTCCATTTGACAGAAAGCCGGAAGTCCATATCGGGTACAAAGTGGGCCACA 2012
 Qy 641 GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 Db 2013 GAAGTACGGGGGCTTCTTCTGAGCTCTGTGTCAAGTTCACCTGATCTTAATGATGTG 2072
 Qy 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 2073 AATGACAAACCTCCAGGCTAGCAAGGACTACACGGGCTGTTCTTCTTCGCAATCCCTC 2132
 Qy 681 SerIleProGlySerLeuIlePheGlnAlaThrAspAspAspGlnIleLeuPheArgLys 700
 Db 2133 AGTGACACCTGGAGTCTCATTTTCAGGCTACATGATGATGATACGACTTAATTTCCGGGT 2192
 Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspThrProGlyValSerLys 720
 Db 2193 CCCCATTTTACATTTTCCCTCGGAGTGAAGCTTCAAAAACACATGGGGAAGTTTCCAAA 2252
 Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGlnGluAlaIleArg 740
 Db 2253 ATCAATGTAATCATGCGGACGCTGTCTACACAGGACACAGAGTTTGAGAGAGAGGATAT 2312
 Qy 741 ValValLeuIleArgIleAsnAspGlyValArgProProLeuGlnGlyIleValSerLeu 760
 Db 2313 GTCTGTCTATGATCGCATCATGATGAGGGGTGCGCCACCTTGAAGGCAATTTGTTCTTTA 2372

Qy 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
 Db 2373 CCAATTACATTCCTGCACTTGTGTGAAAGAGATTTTCCGGCAGCAGGTCAACCAACT 2432
 Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2433 GGGATACCCCACTGTGGGCAATGCGAGTGTGTAATCTGTGACCACTTCTGTGATTTGT 2492
 Qy 801 IleIleLeuAlaValAlaPheIleArgIleLysLysAspLysGlyLysAspAsnValGlu 820
 Db 2493 ATATTTTACAGTGTGTTTATCCCATTAAGAGATTAAGCAAAAGATTAATGTTGAA 2552
 Qy 821 SerAlaGlnAlaSerGluValLysProLeuArgSer 832
 Db 2553 AGTCTCAACATCTGAAGTCAAACTCTGAGAAC 2588

RESULT 8
 US-10-025-380-1086
 ; Sequence 1086, Application US/10025380
 ; Publication No. US20020182191A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Secrist, Heather
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Meagher, Madeleine Joy
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Jiang, Yuguang
 ; APPLICANT: Smith, Carole L.
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Clapper, Jonathan D.
 ; APPLICANT: Skelky, Yashir A. W.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Vedvick Thomas S.
 ; APPLICANT: Carter, Darrick
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
 ; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 210121.471C14
 ; CURRENT APPLICATION NUMBER: US/10/025,380
 ; NUMBER OF SEQ ID NOS: 1129
 ; SOFTWARE: PasteSeq for Windows Version 4.0
 ; SEQ ID NO 1086
 ; LENGTH: 2877
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-025-380-1086

Alignment Scores:
 Pred. No.: 0 Length: 2877
 Score: 4208.00 Matches: 813
 Percent Similarity: 99.51% Conservative: 1
 Best Local Similarity: 99.39% Mismatches: 2
 Query Match: 97.38% Indels: 2
 DB: 9 Gaps: 1

US-10-025-380-1081 (1-832) x US-10-025-380-1086 (1-2877)

Qy 17 LeuAlaThrGly-----TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMet 34
 Db 421 CTGGTACGGGATCCGGGCGCCGACAAAGAGGAGATTGAGGACCCCTGAAACCCATG 480
 Qy 35 ThrPheSerIleTyrGlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAsn 54
 Db 481 ACATTTTCAATTTATGAAGCCCAAGAACCGATCAATTAATTCAGTTTAAGCCAT 540
 Qy 55 ProProAlaValIleThrPheGluLeuThrGlyGluThrAspAsnIlePheValIleGluArg 74
 Db 541 CCTCTGCTGTGATCTTTGAACTTAATCTGGGAGACAGACAAATATTTGTGATAGAACGG 600

QY 75 GluGluLeuLeuTyrTyrTrpAsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeu 94
 DB 601 GAGGAGCTTCGTAATTCACACAGACCTTGACGGGAAACAGATCTTACTACATCTC 660
 QY 95 GlnValAlaAlaLeuAspAlaAsnGlyIleIleValGluGlyProValProIleThrIle 114
 DB 661 CAGGTTGACACCCCTCGACGGCTAATGGAATTATGTGAGGGGTCCAGTCCCTATCACCATA 720
 QY 115 GluValIleAspIleLeuAspAsnArgProThrPheLeuGlnSerIleValIleGluGlySer 134
 DB 721 GAAGTGAAGACATCAACGACCAATCGACCACTTCTCCAGTCCAAAGTACGAAGGCTCA 780
 QY 135 ValArgGlnAsnSerArgProGlyIleValProPheLeuTyrValAsnAlaThrAspLeuAsp 154
 DB 781 GTAAGGCAAGACTCGCCCAAGGAAAGCCCTTCTTGATGTCAATGCCACAGACTGGAT 840
 QY 155 AspProAlaThrProAsnGlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIle 174
 DB 841 GATCCGGCCACTCCCAATGGCCAGCTTATTAACCAATGTGCATCCAGCTTCCCATGATC 900
 QY 175 AsnAsnValMetTyrPheGlnIleAsnAsnIleThrGlyAlaIleSerLeuThrArgGlu 194
 DB 901 AACATGTATATGTAATTCCTTCAGATCAACACAAACGGAGGCCATCTCTTACCCGAGAG 960
 QY 195 GlySerGlnGluLeuAsnProAlaIleAspProSerTyrAsnLeuValIleSerValIle 214
 DB 961 GGATCTCAGGAATGAAATTCCTGCTAAGAAATCTCTTAATATCTGAGTCTCGATAG 1020
 QY 215 AspMetGlyGlyGlnSerGluAsnSerPheSerAspThrThrSerValAspIleIleVal 234
 DB 1021 GACATGGAGGCCAGAGTGAATTCCTTCAGTATACCAATCTGTGATATCATATAGTG 1080
 QY 235 ThrGluAsnIleTyrPheAlaProIleProValGluMetValGluAsnSerThrAspPro 254
 DB 1081 ACAGAGATATTTGGAAAGCACCAAACTGTGAGATGGGGAACCTCAACTGATCTCT 1140
 QY 255 HisProIleValIleThrGlnValArgTrpAsnAspProGlyAlaGlnIleTyrSerLeuVal 274
 DB 1141 CACCCATCAAAATCACTCAGGTGCGGTGAAATATCCGGTGCACATATTTCTTAGTT 1200
 QY 275 AspIleGluIleAspProArgPheProPheSerIleAspGlnGluGlyAspIleTyrVal 294
 DB 1201 GACAAAGAGAGCGCCAAAGATTCCTCAATTTCAATTTGACGAAAGGATATTATTCGTG 1260
 QY 295 ThrGlnProIleAspArgGluGluIleAspAlaTyrValPheTyrAlaValAlaIleAsp 314
 DB 1261 ACTCAGCCCTTGGACCGAGAAAGATGCATATGTTTTTATGCGATGGCAAAAGAT 1320
 QY 315 GluTyrGlyIleAspProLeuSerTyrProLeuGlnIleHisValIleValIleAspIleAsn 334
 DB 1321 GAGTACGGAAACCACTTTCATATCCGCTGAAATTCATGTAAAGTTAAAGATATTAAT 1380
 QY 335 AspAsnProProThrCysProSerProValThrValPheGluValGlnGluAsnGluArg 354
 DB 1381 GATATATCACTCAATGTCGCTCAACGATACCGATATTGAGTCCAGAGAAATGAAGA 1440
 QY 355 LeuGluIleAsnSerIleGlyThrLeuThrAlaHisAspArgGluGluAsnThrAlaAsn 374
 DB 1441 CTGGGTAAACGATACGGGACCTTACGCAATGACAGGGATGAAGAAATACGCCAAC 1500
 QY 375 SerPheLeuAsnTyrArgIleValGluGlnThrProIleAspProMetAspGlyLeuPhe 394
 DB 1501 AGTTTCTAACTACAGGATTTGGAGCAACTCCCAAACTTCCCATGGATGGACTCTTC 1560
 QY 395 LeuIleGlnThrTyrAlaGlyMetLeuGlnLeuAlaIleGlnSerLeuIleValIleAsp 414
 DB 1561 CTATATCAAACTATGCTGGAATGTACAGTTAGTAAACAGTCTTGAAGAACACAGAT 1620
 QY 415 ThrProGlnTyrAsnLeuThrIleGluValSerAspIleAspPheIleThrLeuCysPhe 434
 DB 1621 ACTCTCAGTACCACTTAACGATAGAGGTGTCTGCAAAAGATTTCAAGAACCTTGTGTTT 1680
 QY 435 ValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluIleSerAspTyr 454
 DB 1681 GTGCAAATCAAGCTTATGATATCAATGATCAGATCCCATCTTTGAAAAATCAGATTAT 1740
 QY 455 GlyAsnLeuThrLeuAlaGluAspThrAsnIleGlySerThrIleLeuThrIleGlnAla 474
 DB 1741 GGAAACCTGACTTGTCTGAAGACAAACATTTGGGTCCACCATTTAACTCAACCGAGCC 1800
 QY 475 ThrAspAlaAspGluProPheThrArgIleSerSerIleLeuTyrHisIleIleIleValGly 494
 DB 1801 ACTGATGCTGAATGAGCAATTTACTGGAAGTTCTAAATCTGTATCATATCAATAAAGGA 1860
 QY 495 AspSerGluGlyArgLeuGlyValAspThrAspProHisThrAsnThrGlyTyrValIle 514
 DB 1861 GACAGTGAAGGAGCGCTGGGGGTTCACACAGATCCCATACCAACACCGGATATGCATA 1920
 QY 515 IleIleValProLeuAspPheGluThrAlaAlaValSerAsnIleValPheIleValGlu 534
 DB 1921 ATTTAAAGCCCTTATATTTTGAACAGACGCTTTCCAACTGTGTTCAAAGCAGA 1980
 QY 535 AsnProGluProLeuValPheGlyValIleTyrAsnAlaSerSerPheAlaIlePheThr 554
 DB 1981 AATCTGAGCCTTATGTTGTGTGTGAGTACAAATGCAAGTCTTTTCCAGATTCAG 2040
 QY 555 LeuIleValThrAspValAsnGluAlaProGlnPheSerGlnHisValPheGlnAlaIle 574
 DB 2041 CTATATTGACAGATGTAATGAAGACCTCAATTTCCCAACAGTATTCMAAGCGAAA 2100
 QY 575 ValSerGluAspValAlaIleGlyThrIleValIleValAsnValThrAlaIleAspProGlu 594
 DB 2101 GTCAGTGAAGATGTACTATGACCTAAAGTGGGAAATGTGCTCCAAAGATCCAGAA 2160
 QY 595 GlyLeuAspIleSerTyrSerLeuAspArgIleAspThrArgGlyTyrPheIleIleAspHis 614
 DB 2161 GGTCTGACATTAAGCTATTCATCTGAGGAGACACAAAGCTTGGCTTAAATTAAGCCAC 2220
 QY 615 ValThrGlyGluIlePheSerValAlaProLeuAspArgGluAlaGlySerProTyrArg 634
 DB 2221 GTGACTGGTGAAGTCTTATGATGCTCCATTTGACAGAGAGCCGGAAGTCCATATCCG 2280
 QY 635 ValGlnValAlaThrGluValGlyIleSerSerLeuSerSerValSerGluPheHis 654
 DB 2281 GTCAAGTGTGTGCCACAGAAAGTGAAGGGGTCTTCTTAAGCTGTGTGACAGTTCCAC 2340
 QY 655 LeuIleLeuMetAspValAsnAspAsnProProArgLeuAlaIleAspTyrThrGlyLeu 674
 DB 2341 CTATCTCTTAATGATGTAATGACAACTCCAGGCTACCAAGACATACAGGGCTTG 2400
 QY 675 PhePheCysHisProLeuSerAlaProGlySerLeuIlePheGluAlaThrAspAsp 694
 DB 2401 TTCTTTCGCAATCCCTCAGTGCACCTGGAAGTCTCATTTTGAAGCTACTGATGATAT 2460
 QY 695 GlnHisLeuPheArgGlyProHisIlePheThrPheSerLeuGlySerGlySerLeuGlnAsn 714
 DB 2461 CAGACCTTATTTGGGGGTCCCATTTTACATTTTCCCTGGCAGTGAAGCTTACAAAC 2520
 QY 715 AspTrpGluValSerIleValIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAsp 734
 DB 2521 GACTGGGAAGTTTCCAAATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2580
 QY 735 PheGluGluValArgAlaTyrValValLeuIleArgIleAsnAspGlyIleArgProProLeu 754
 DB 2581 TTTGAGAGAGGGCGATATGCTGCTGATCCGATCAATATGATGGGGGTCCGCCACCTTG 2640
 QY 755 GluGlyIleValSerLeuProValThrPheCysSerCysValGluGlySerCysPheArg 774
 DB 2641 GAAAGCATTTGTTCTTATCAAGATTCATCTGCAAGTGTGTGGAAGGAAGATTTTTCGG 2700
 QY 775 ProAlaGlyHisGlnThrGlyIleProThrValIleGlyMetAlaValGlyIleLeuLeuThr 794
 DB 2701 CAGCAGGTACCAAGCTGGGATACCACTGTGGCATGGCAATGGTATATCTGCTGACC 2760
 QY 795 ThrLeuLeuValIleGlyIleIleLeuAlaValPheIleArgIleIleValIleValIleVal 814

```

Db      2761 ACCCTTGCTGATGATGATTAATTTAGACAGTGTGTTTATCCGCATTAAGAGATPAA 2820
Qy      815 G1yLysAspAsnVal1G1uSer1aG1nAlaSer1uVal1LysProLeuArgSer 832
      2821 GGCAAGATATATGTTGAAAGTGCTCAAGCATCTGAAGTCAAACTCTGAGAGC 2874

RESULT 9
US-09-922-217-1086
; Sequence 1086, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secret, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Mesgher, Madeleine Joy
; APPLICANT: Stoik, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922.217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ. ID NOS: 1124
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1086
; LENGTH: 2877
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-217-1086

Alignment Scores:
Pred. No.: 0 Length: 2877
Score: 4208.00 Matches: 813
Percent Similarity: 99.51% Conservative: 1
Best Local Similarity: 99.39% Mismatches: 2
Query Match: 97.38% Indels: 2
DB: 10 Gaps: 1

US-10-025-380-1081 (1-832) x US-09-922-217-1086 (1-2877)
Qy      17 LeuAlaThrGly-----TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMet 34
      421 CTGCGTACCGGATCCGCGCGCCGACAAAGAGGGGAAGTTTGTGACCCCTGAAACCCCATG 480
Db      35 ThrPheSer1LeuTyrGluGlyGlnGluProSerGln1LeuPheGlnPheLysAlaAsn 54
      481 ACATTTCTATTAATGAAGGCAAGAAACGAGTCAATTAATTCAGTTTAAGGCAAT 540
Qy      55 ProPro1aValThrPheGluLeuThrGlyGluThrAspAsn1LeuPheVal11LeuArg 74
      541 CTCCTGCTGCTGACTTTTGAACCTACCTGCGGAGACACAAATATTTGTGATGAACGG 600
Db      75 GlnGlyLeuLeuTyrTyrAsnArgAlaLeuAspArgGluThrArgSerThrAsnLeu 94
      601 GAGGAGCTTCTGTATTAACAACAGAGCTTGGACAGGAAACAAGATCTACTCAATCTC 660
Qy      95 GluValAlaAlaLeuAspAlaAsnGlyLeuValGlnGlyProValPro1LeuThr1Le 114
      661 CAGGTTGACCCCTCGACGCGCTAATGAAATTAATGAGGAGGTCCAGTCCCTATCACATA 720
Db      115 GluValLysAsp1LeuAspAsnArgProThrPheLeuGlnSer1LysTyrGluGlySer 134
      721 GAAGTGAAGACATCAACGAAATCGAACCCAGTCTTCCAGTCAAGTAAAGTACGAAGCTCA 780
Qy      135 ValArgLysAsnSerArgProGlyLysProPheLeuTyrValAsnAlaThrAspLeuAsp 154
      781 GTAAGGCAAACTCTCGCCAGAAAGCCCTTCTTGTATGTCAATGCCACAGACTGGAT 840

```

```

Qy      155 AspProAlaThrProAsnGlyGlnLeuTyrTyrGln1LeuVal11LeuGlnLeuProMet1Le 174
      841 GATCCGGCCACTCCCAATGCGCAGCTTTATTAACAGATTGTCAAGCTTCCCATATC 900
Db      175 AsnAsnVal1MetTyrPheGln11LeuAsnLysThrGlyAla11LeuSerLeuThrArgGlu 194
      901 AACATATGTATGATCTTTCAGATCAACCAACAAACGGAGCCACTCTCTTACCCGAGAG 960
Qy      195 GlySerGln1LeuLeuAsnProAlaLysAsnProSerTyrAsnLeuVal11LeuSerValLys 214
      961 GGATCTCAGGAATGGAATCTGCTAAGAAATCTTCTATATCTGTGATCTCAGTGAAG 1020
Db      215 AspMetGlyGlyGlnSerGluAsnSerPheSerAspThrThrSerValAsp1LeuVal 234
      1021 GACATGGAGGCGCAGATGGAATTCCTTACGATACCACTGTGGATATCATATGTC 1080
Qy      235 ThrGluAsn11LeuTyrLysAlaProLysProValGluMetValGluAsnSerThrAspPro 254
      1081 ACAGAGAAATATTGAAAGCACCMAAACCCTGTGAGATGGTGGMAAATCAACTGATCCT 1140
Db      255 HisPro1LeuLys1LeuThrGlnValArgTyrAsnAspProGlyValGlnTyrSerLeuVal 274
      1141 CACCCCATCAAAATCACTCAGGTGCGGTGGAATGATCCCGTGCACAAATATCTTAAGT 1200
Qy      275 AspLysGluLysLeuProArgPheProPheSer1LeuAspGlnGluGlyAsp1LeuTyrVal 294
      1201 GACAAAGAAAGCTGCGCAATTCCTCAATTTCAATTAAGCAGAAAGAGATATTTACGTG 1260
Db      295 ThrGlnProLeuAspArgGluGluLysAspAlaTyrValPheTyrAlaValAlaLysAsp 314
      1261 ACTCAGCCCTTGGACGAGAAAGAAAGATGATATGTTTTTATGCAAGTGGCAAGGAT 1320
Qy      315 GlnTyrGlyLysProLeuSerTyrProLeuGln11LeuValLysValLysAsp1LeuAsn 334
      1321 GAGTACGAAAGAACCATTTCTATATCCGCTGGAATCATGTAAAGTAAAGATATTAAAT 1380
Db      335 AspAsnProProThrCysProSerProValThrValPheGlnValGlnGluAsnGluArg 354
      1381 GATTAATCACTTCAATGATCCGTCAACAGTAACTGTTTGAAGTCCAGGAAATGAACGA 1440
Qy      355 LeuGlyAsnSer1LeuGlyThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsn 374
      1441 CTGGGTAAACGATACCGGACCCCTTACTGCACATGACAGGAGTGAAGAAATATCTGCCAAC 1500
Db      375 SerPheLeuAsnTyrArg1LeuValGlnGlnThrProLysLeuProMetAspGlyLeuPhe 394
      1501 AGTTTCTTAAACTTACAGGATGTGGAGCAAACTCCCAAACTTCCCATGATGACTCTTTC 1560
Qy      395 Leu11LeuGlnThrTyrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAsp 414
      1561 CTAAATCAAACTTATCTGAAATGTTTACAGTTAGCTTAAACAGTCTTGAAGAAACAGAT 1620
Db      415 ThrProGlnTyrAsnLeuThr11LeuValSerAspLysAspPheLysThrLeuCysPhe 434
      1621 ACTCCTCAGTACAACTTAACGATAGAGGTCTCGACAAAGATTTCAGAGCCCTTGTGTTT 1680
Qy      435 ValGln1LeuAsnVal11LeuAsp1LeuAsnAspGln1LeuPro1LeuPheGlnLysSerAspTyr 454
      1681 GTCCAAATCAACGTTATGATATCAATGATCAAGATCCCACTTTGAAAAATACAGATTAT 1740
Db      455 GlyAsnLeuThrLeuAlaGluAspThrAsn11LeuSerThr1LeuThr1LeuGlnAla 474
      1741 GGAACCTGACTTCTTCTGTAAGACAAACATTTGGGTCCACCAATCTTAACATCCAGGCC 1800
Qy      475 ThrAspAlaAspGluProPheThrGlySerSerLys1LeuTyrHis11LeuLysGly 494
      1801 ACTGATGCTGATGAGCACTTACTGAGAGTTCTAAATTTCTGATCATATCAATAAGGGA 1860
Db      495 AspSerGlnGlyArgLeuGlyValAspThrAspProHisThrAsnThrGlyTyrVal11Le 514
      1861 GACAGTGAAGGAGCGCTGGGGGTGACACAGATCCCATATACCAACCGGATATGTGCATA 1920

```

QY 515 ILELYSLVPRLEUASPPHEGLUTHRALALAVASERANILEVALPHELYSALAGLU 534
 DB 1921 ATTTAAAGGCTCTTGTATTTGAAACAGACGCTTTTCCAACTGTTGCTTCAAGCGAA 1980
 QY 535 ASNPRLUPLROLEUVALPHEGLYVALYSLYTRNANLASERSEPRHEALALYSPHETHR 554
 DB 1981 AATCTAGGCTCTAGTGTGTGGGTGAGTACAAATCAATGCTTTTGGCAAGTTCACG 2040
 QY 555 LEUULEVALTHRAPVALAENGUVALAPROGLNPHESERGINHISVALPHEGLNALALYS 574
 DB 2041 CTTATTTGACAGATGGAATGGAAGACCTCAATTTTCCAAACGATTCAGCGAA 2100
 QY 575 VALSERGLUPSPVALALALEGLYTHRYEVALGLYASNAVLTTHRALYASPPROGLU 594
 DB 2101 GTCACTGAGATGATGCTATGAGCACTTAAGTGGGCAATGTGACTGCCAAGATCCGAA 2160
 QY 595 GLYLEUASPILESERTYRSELEUARGLYASPTHARGLYTPLEULYSLLEASPHIS 614
 DB 2161 GGCTGACATTAAGCTATTCTAGAGGGAGACACAAAGAGGTGGCTTAAATGACAC 2220
 QY 615 VALTHRGLYGLULIEPHESERVALALAPROLEUASPPROGLUAGLYSERPROTYRARG 634
 DB 2221 GTGACTGTGAGATCTTATGATGCTCCATTTGACAGAGAGCCGGAAGTCCATATCGG 2280
 QY 635 VALGLNAVJALVALATHRGUVALGLYLSERSELEUSERSERVALSERGLUPHEHIS 654
 DB 2281 GTACAGATGATGACCAAGAGTGGGGGCTTCTTAACTTGTGTGAGATTCAC 2340
 QY 655 LEUULEUWETASPPVALASNPASNPBROBROHLEUVALALYASPTYRTHGLYLEU 674
 DB 2341 CTGATCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
 QY 675 PHEPHECYHISAPROLEUSERVALAPROGLYSERLEULEPHEGLUALATHRAPASAPSP 694
 DB 2401 TTTCTTCCGATCCCTCCAGACCTGGAAGTCTCATTTTCCAGGCTACGATGATGAT 2460
 QY 695 GLNHSLEUPHEARGLYPROHISPHETHRPHESERLEUGLYSERGLYSERLEUGLNASN 714
 DB 2461 CAGACCTTATTTCCGGGCTCCCATTTTTCATTTTCCCTCCGCACTGGAAGCTTCAAAAC 2520
 QY 715 ASPTRGLUVALSERLYSLLEASNGLYTHRHISALARGLEUSERTHRARGHISTHRASP 734
 DB 2521 GACTGGGAGATTTCCAAATCAATGATGATGATGATGATGATGATGATGATGATGAT 2580
 QY 735 PHEGLUGUARGALATYRVALVALLEULEHARGLYLEASNAPGLYARGPROPROLEU 754
 DB 2581 TTTGAGAGAGAGGGGTATGCTGCTGATCCGATCAATGATGATGATGATGATGATGAT 2640
 QY 755 GLUGLYLEVALSERLEUPROVALTHRPHESERCYSEVALGLUGLYSERCYSPHEARG 774
 DB 2641 GAAGGCAATGTTCTTTTACAGATTACATTTCTGCACTGTGTGGAAGAGATTTTCCG 2700
 QY 775 PROHLAGLYHISINTHGLYLSERPROTHRVAGLYMETALVALGLYLEULEUETHR 794
 DB 2701 CCACAGGATCCACAGACTGGATACCCACTGTGGGCAATGATGATGATGATGATGAT 2760
 QY 795 THRILEUVALILEGLYILEILEUVALVALPHELEHARGLYLEULYSLYASPLYS 814
 DB 2761 ACCCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
 QY 815 GLYLYSAPPAENVALGLUSERVALGLNALASERGLUVALYSPROLEUARGSER 832
 DB 2821 GGCAAGATATATGTTGAAAGTCTCAAGCATCTGAAGTCAAACTCTGAGAGAC 2874

RESULT 10
 US-09-833-263-1086
 ; Sequence 1086, Application US/09833263
 ; Patent No. US20020110547A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Clapper, Jonathan D.
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Meagher, Madeleine J.

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 210121.471C12
 ; CURRENT APPLICATION NUMBER: US/09/833,263
 ; CURRENT FILING DATE: 2001-04-10
 ; NUMBER OF SEQ ID NOS: 1093
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1086
 ; LENGTH: 2877
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (1)...(2877)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-833-263-1086

Alignment Scores:
 Pred. No.: 0
 Score: 4208.00
 Percent Similarity: 99.51%
 Best Local Similarity: 99.39%
 Query Match: 97.38%
 DB: 10
 Gaps: 1

US-10-025-380-1081 (1-832) x US-09-833-263-1086 (1-2877)

QY 17 LEUVALATHGLY-----TYRGLYGLINGLYLYSPHESERGLYPROLEUYSPPROMET 34
 DB 421 CTGCTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 480
 QY 35 THRPHESERLIETRYGLUGLYGLINGLYUPROSEGLNILELEPHEGLNPHELYSALASN 54
 DB 481 ACATTTTCTATTTATGAGGCCAAGAACGAGCAATATATATTCAGTTTAAGGCCAAT 540
 QY 55 PROPRALAVATHRPHEGLEUETHRGLYGLUTHRAPASNTLEPHEVALILEGLUARG 74
 DB 541 CCTCGCTGTGACTTTTGAATCACTGAGGGAGACAGACAACTATTTGATGAGAACGG 600
 QY 75 GLUGLYLEUETHRYRTHRAPASNARGALALEUASPPARGLYTHRPSERTRHISANLEU 94
 DB 601 GAGGACCTTCTGATTAACAACAGAGCTTGGACAGGAAACAGATCTACTCAATCTC 660
 QY 95 GLNVALALALEUASPPALASNGLYILELEVALGLUGLYPROVALPROILETHRTILE 114
 DB 661 CAGCTTCAAGCTTGAAGCGTATGATGATGATGATGATGATGATGATGATGATGAT 720
 QY 115 GLUVALYASPILEASNPASNPARGPROTHRPHLEUGLINSERYTYRGLUGLYSER 134
 DB 721 GAAGTGAAGACATCAACAGCAATCGAACCCCTTCTCCAGTCAAAAGTACGAAGGCTCA 780
 QY 135 VALARGLEASNPASNPARGPROGLYLYSPROPHLEUETHRVVALASNPALATHRAPLEUASP 154
 DB 781 GTAAGGAGAACTCTGCCCGAAGAACCTTTTGTATGTCATGTCACAGACCTGGAT 840
 QY 155 ASPPROLATHRPROASNGLYGLNLEUETHRYTYRGLNILEVALILEGLINLEUPROMETILE 174
 DB 841 GATCCGGCCACTCCCAATGAGCCGCTTATTAACAATGTCTACAGCTTCCCATGATC 900
 QY 175 ASNPASNAVLETYRPHENGLILEASNAENLYETHRGLVALILESERLEUETHRARGGLU 194
 DB 901 AACCAATGTATGATCTTTCAGATCAACAAACGAGGAGCCATCTCTTACCCGAGAG 960
 QY 195 GLYSERGLNLEUENPVALALYASNPASNPARGPROSERTRYRASNLEUVALILESERVALYS 214
 DB 961 GATCTCAGGAATGATCTCTGTAAGAACCTTCTCATATATGTGGATCTCAGTGAAG 1020
 QY 215 ASPMERCGLYGLINSERGLUASNPASNPASNPASNPASNPASNPASNPASNPASNPASNP 234
 DB 1021 GACATGGAGGCCAGATGAGATTTCTTCAAGATACCAATCTGTGGATACATATGATG 1080
 QY 235 THRGUASNTLETRPYVALAPROLYSPROVALGLUWETVALGUASNPASNPASNPASNP 254

Db 1081 ACAGAGATATTTGGAAAGCAACCAAACTGTGAGATGTGGAAAACTCACTGATCT 1140
 QY HisProileuyllethrglnValArgTrrAsnAspProglValaglnTyrSerLeuVal 274
 Db 1141 CAACCCATCAAAATCATCTCAGTGCCTGGATGATATCCCGGTGCACATATCTTGGT 1200
 QY AspIyglulysleuProArgpheProPheSerlleaspglnglulysAspIleTyrVal 294
 Db 1201 GACAAAGAGAAAGCTGCAGATTCCTCAATTCATGACAGAGAGATATTTACGTG 1260
 QY ThrGlnProleuAspArgglulglulysAspAlaTyrValPheTyrAlaValAlaIysAsp 314
 Db 1261 ACTAGCCCTTGGACCGAGAAAGAAAGATGCATATGTTTTTATGAGATGGCAAGAGAT 1320
 QY GlulTyrGlyLysProleuSerTyrProleuGlnlleIysValLysValLysAspIleAsn 334
 Db 1321 GAGACGGAAGAAACACCTTTCATATCCCTGGAAATTCATGTAAAGTTAAAGATATTAAT 1380
 QY AspAsnProProthTyrCysProSerProValThrValPheglulValGlnIyAsnGluArg 354
 Db 1381 GATATATCACCCTACATGTCGTGCACAGTAAACGTATTTGAGGTCCAGAGAAATGACGA 1440
 QY 355 LeuGlyAsnSerlleglTyrThrleuThrAlaHisAspArgAspGluGluAsnThrAlaAsn 374
 Db 1441 CTGGGTAAACAGTATCGGACCTTATGCAATGACATGACAGGATGAAAGAAATACGCGAAC 1500
 QY SerPheleuAsnTyrArgIleValGlnGlnThrProLysLeuProMetAspGlyLeuPhe 394
 Db 1501 AGTTTCTTAAACATACAGAGATTTGTGACAAATCCCAACTTCCCATGATGACATCTTC 1560
 QY 395 LeuIleGlnThrTyrAlaGlyMetLeuGlnleuAlaLysGlnSerLeuLysGlnAsp 414
 Db 1561 CTAATCCAAACATATGCTGAGATGTACAGTACCTAAACAGTCTTGAAGAGCAAGAT 1620
 QY 415 ThrProGlnTyrAsnleuThrIleGluValSerAspLysAspPheLysThrLeuCysPhe 434
 Db 1621 ACTCTCAGTAAACATTAACAGATAGAGGTGTGACAAAGATTTCAAGACCTTTGTTT 1680
 QY 435 ValGlnlleAsnVallleAspIleAsnAspGlnleProIlePheglulLysSerAspTyr 454
 Db 1681 GTCCAAATCAACGTATATGATATCAATGATCAATCCCATCTTTGAAAATCAAGATTAAT 1740
 QY 455 GlyAsnleuThrleuAlaGluAspThrAsnIleGlySerThrIleleuThrIleGlnAla 474
 Db 1741 GGAACCTGACTCTGTGTGAAGACACAACATTTGGGTCCACATCTTAACATCCAGGCC 1800
 QY 475 ThrAspAlaAspGluProPheThrGlySerSerIyslleuTyrHisIlelleLysGly 494
 Db 1801 ACTGATGCTGATAGCCATTTACTGGGAGTTCTTAAATTCGTATCATATCATTAAGGGA 1860
 QY 495 AspSerGluGlyArgleuGlyValAspThrAspProHisThrAsnThrGlyTyrValIle 514
 Db 1861 GACAGTGAAGGAGCCCTGGGGGTTTGACAGATCCCATACCAACACCGGATATGCTATA 1920
 QY 515 lleLysLysProleuAspPheglulThrAlaValSerAsnIleValPheLysAlaGlu 534
 Db 1921 ATTAAGAAAGCCTCTGATTTTGAAGACAGCGCTTTCCAAATGTGTTCAAGACGAA 1980
 QY 535 AsnProGluProleuValPheglulValLysTyrAsnAlaSerSerPheAlaLysPheThr 554
 Db 1981 AATCCGAGGCTCTAGAGTTGGTGAAGTAACTCAAGTCTTTTGGCAAGTTCACG 2040
 QY 555 LeuIleValThrAspValAsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLys 574
 Db 2041 CTTATTTGTGACATGTGAATGAAGCACTCAATTTTCCCAACAGATATCCAGGCAAA 2100
 QY 575 ValSerGluAspValAlaIleGlyThrLysValGlyAsnValThrAlaLysAspProGlu 594
 Db 2101 GTCAGTGAAGGATGATGATAGGCACTAAAGTGGGCAATGTGACTGCCAAGATCCAGAA 2160
 QY 595 GlyLeuAspIleSerTyrSerleuArgGlyAspThrArgGlyThrleuLyslleAspHis 614
 Db 2161 GGTCTGACATTAAGCTATTCACTGAAGGGAGACACAGAGAGTTGGCTTAAATTTGACAC 2220

QY 615 ValThrGlyGlnlePheSerValAlaProleuAspArgGluAlaGlySerProTyrArg 634
 Db 2221 GTGACGTGAGATCTTTAGTGTGCTCCATTGACAGAGAACCGGAAGTCAATTCG 2280
 QY 635 ValGlnValAlaAlaThrGluValGlyLysSerSerleuSerValSerGluPheHis 654
 Db 2281 GTACAGTGGTGCACAGAAAGTGAAGGGGCTTCTTAAAGCTGTGTGACAGTTCCAC 2340
 QY 655 LeuIleleuSerAspValAsnAspAsnProProArgleuAlaLysAspTyrThrGlyLeu 674
 Db 2341 CTGATCTTATGATGTGAATGACAAACCTCCAGGCTGACCAAGAGACTACAGGGCTTG 2400
 QY 675 PhePheCysHisProleuSerAlaProGlySerLeuIlePheGlnAlaThrAspAsp 694
 Db 2401 TTCTTGTCCATCTCCCTCAGTGCACCTGGAAGCTCATTTTGAAGCTACTGATGATGAT 2460
 QY 695 GlnHisleuPheArgGlyProHisPheThrPheSerleuGlySerGlySerleuGlnAsn 714
 Db 2461 CAGACCTTATTTGCGGGTCCCATTTTACATTTTCCCTCGGAGTGAAGCTTACAAAC 2520
 QY 715 AspTrrGluValSerIyslleAsnGlyThrHisAlaArgleuSerThrArgHisThrAsp 734
 Db 2521 GACTGGAAAGTTCCAAATCAATGATGACTCAATGCCGACTGTACAGGACACAGAC 2580
 QY 735 PheGlnGluArgAlaTyrValValleuIleArglleAsnAspGlyGlyArgProProleu 754
 Db 2581 TTGAGAGAGAGGCGATATGCTGTTGATCCGATCAATGATGGGGGTGGCCACCTTG 2640
 QY 755 GlulGlylleValSerleuProValThrPheCysSerCysValGlnGlySerCysPheArg 774
 Db 2641 GAAGCATTTGTTCTTTACCATGATCATCTGCAAGTTGTGGAGAAAGATTTGTTCCGG 2700
 QY 775 ProAlaGlyHisGlnThrGlyLysProThrValGlyMetAlaValGlylleleuThr 794
 Db 2701 CCAGCAGGTACACAGCTGGATACCCACTGTGGCATGGCATGGTGTATACGTGACC 2760
 QY 795 ThrleuLeuVallleGlyllelleleuAlaValPheIleArglleLysLysAspLys 814
 Db 2761 ACCCTTCTGTGATGTGATTAATTTTGAAGATTTGTATTCGCTAAAGAAAGATTA 2820
 QY 815 GlyLysAspAsnValGluSerAlaGlnAlaSerGluValLysProleuArgSer 832
 Db 2821 GGCAAGATTAATGTTGAAGTGTCAAGCATGTGAAGTCAACCTCTGAAGAC 2874

RESULT 11
 US-10-063-547-97
 ; Sequence 97, Application US/10063547
 ; Publication No. US20020182638A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Matanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3230R1C1
 ; CURRENT APPLICATION NUMBER: US/10/063,547
 ; PRIORITY FILING DATE: 2002-05-02
 ; NUMBER OF SEQ ID NOS: 170
 ; SEQ ID NO 97
 ; LENGTH: 2848
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 ; US-10-063-547-97

Alignment Scores:

Pred. No.: 3.24e-89 Length: 2848
 Score: 908.50 Matches: 237
 Percent Similarity: 50.51% Conservative: 159
 Best Local Similarity: 30.23% Mismatches: 335
 Query Match: 21.03% Indels: 53
 DB: 9 Gaps: 21

US-10-025-380-1081 (1-832) x US-10-063-547-97 (1-2848)

QY 62 LeuThrglyglu-----ThraSpasnIlePheValIleGluThrglyu---Gly 76
 DB 296 CTGTCAGGGGACTCAGGCAAGCACTGAGGGCCCATTTGCTATGAGTCAAGATTCTGGC 355
 QY 77 LeuLeuTyrrTaAsnArgAlaLeuAspArgGluThraArgSerThrHisAsnLeuGlnVal 96
 DB 356 TTCCTGCTGGTACCAAGGCCCCCTGACCCAGAGAGACGACGAGTACCACTCAGAGTC 415
 QY 97 AlaAlaLeuAspAlaAsnGlyIleIleValGluGlyProValProIleThrIleGluVal 116
 DB 416 ACCCTGGAGATGACAGATGACATGCTTGTGGGGTCCACAGCCTGTGCTTGTCAGCTG 475
 QY 117 LysAspIleAsnAspAsnArgProThrPheLeuGlnSerLysTyrrGluGlySerValArg 136
 DB 476 AAGGATGAGATGACCAAGTGCCTCATTTCTCAAGCCATCTACAGAGCTCGGCTAGC 535
 QY 137 GlnAsnSerArgProGlyLysProPheLeuTyrrValAsnAlaThrAspLeuAspAspPro 156
 DB 536 CGGGGTACAGAGCCCTGGATCCCTCTCTCTCTGAGGCTTCAAGCCGAGATGAGCA 595
 QY 157 AlaThrProAsnGlyGlnLeuTyrrGlnIleValIleGlnLeuProMetIleAsnAsn 176
 DB 596 GGCAAGCAAGCTGGATCTTCGATCCATCCAGCCAGGCTCAGCCAGCCAGCTTCC 655
 QY 177 ValMetTyrrPheGlnIleAsnAsnLysThrGlyValIleSerLeuThrArgGluGlySer 196
 DB 656 CCAGACATGTTCCAGCTGAGCTCGGCTGGGGCTGCGCTCAGCCCAAGAGGAGC 715
 QY 197 GlnGluLeuAsnProAlaLysAsnProSerTyrrAsnLeuValIleSerValLysAspMet 216
 DB 716 ACCAGCCTTGACCAAGCCCTGAGAGACCTACCAAGCTTGTGTCAGAGTCAAGATG 775
 QY 217 GlyGlyGlnSerGluAsnSerPheSerAspThrThrSerValAspIleIleValThrGlu 236
 DB 776 GGTACCAAGGCC--TCAGGCCACCAAGGCCACTGCACCTGGAAGTCTCATATGAG 832
 QY 237 AsnIleTyrrPheValProLysProValGluMetValGluAsnSerThrAspProHisPro 256
 DB 833 AGCACCTGGGTGCTCTAGAGCTATCCACTGCAAGAGATCTCAAGTCTTATACCG 892
 QY 257 IleLysIleThrGlnValArgTyrAsnAspProGlyAlaGlnTyrrSerLeuValAspLys 276
 DB 893 CACCAACATGCCCCAGGTACCTGAGAGTGGGGTATGTGCACTTACCTG----- 943
 QY 277 GlnLysLeuProArgPheProPheSerIleAspGlnGluLysAspIleTyrrValThrGln 296
 DB 944 GAGAGCAATCCCGCGGACCTTTGAAGTGAATGAGAGGAAACCTTACGTACAGAGA 1003
 QY 297 ProLeuAspArgGluGlnLysAspAlaTyrrValPheTyrrAlaValAlaLysAspGluTyrr 316
 DB 1004 GAGCTGACAGAGAGCCAGCTGAGTACCTGCTCCAGGTGCGGCTCACAATTCAT 1063
 QY 317 GlyLysProLeuSerTyrrProLeuGlnIleHisValLysValLysAspIleAsnAspAsn 336
 DB 1064 GGGAGAGACTATGGGCCCCCTCTGAGCTGCACGTGCTGAGTGAATGAGATGACAAAC 1123
 QY 337 ProProThrCysProSerProValThrValPheGlnValGlnGluAsnGluArgLeuGly 356
 DB 1124 GTGCTATCTGCTCCCTCCCGTACCCCAAGTCAAGCATCCCTGAGCTCAGTCCACAGGT 1183
 QY 357 AsnSerIleGlyThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPhe 376
 DB 1184 ACTGAAGTACTAGACTGTCTCAGCAGAGATGACAGATGCCCGGCTCCCAATTTCCAC 1243

QY 377 LeuAsnTyrrArgIleValGluGlnThrProLysLeuProMetAspGly---LeuPheLeu 395
 DB 1244 GTTGTGTATAGTCTCTGAGCCCTGAGCCCTGAGAGATGGGGTATGAGGAGAGCTTCCAG 1303
 QY 396 IleGlnThrTyrrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLys---GlnAsp 414
 DB 1304 GTGAGCCCCCACTTCAGAGAGTGTGAGCTGGGGGTCTCCCATCTCCAGAGGCCAGAAC 1363
 QY 415 ThrProGlnTyrrAsnLeuThrIleGluValSerAspLysAsp-----PheLysThrLeu 432
 DB 1364 ATCTGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423
 QY 433 CysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLysSer 452
 DB 1424 TGTGAATGCAAGTCCAGATCCAGATATCAATGATCAAGCCCTGATTCATCACTTCC 1483
 QY 453 AspTyrrGlyAsnLeuThrLeuAlaGluAspThrAsnIleGlySerThrIleLeuThrIle 472
 DB 1484 CAGATTGGGCTTAAAGCTCCCTGAGATGTGAGACCCGGGACTCTGTGGCCATGCTA 1543
 QY 473 GlnAlaThrAspAlaAsp---GluProPheThrGlySerSerLysIleLeu---TyrrHis 490
 DB 1544 ACAGCCATTGATCTGATCTGAGCTCGAGCC-----GCTTCCGCTCATGATTTTGC 1594
 QY 491 IleIleLysGlyAspSerGluGlyArgLeuGlyValAspThrAspProHisThrAsnThr 510
 DB 1595 ATTGAGAGGGGAGACACAGAGAGCATTTTGGCTGATGGAGCA-----GACTCT 1648
 QY 511 GlyTyrrValIleIleLys-----LysProLeuAspPheGluThrAlaAlaValSerAsn 528
 DB 1649 GGGCATGTTAGACTGACCTGCAAGTCAAGCAACTGATTAAGGACACTCCAGTCACTGAG 1708
 QY 529 IleValPheLysAlaGluAsnProGluProLeuValPheGlyValLysTyrrAsnAlaSer 548
 DB 1709 GT 1765
 QY 549 SerPheAlaLysPheThrLeuIleValIleThrAspValAsnGlnAlaProGlnPheSerGln 568
 DB 1766 GCAACCGCCAGCGGTGATCTGATGTGAGAGAGTATGCAACCCCAAGTTGAGACAG 1825
 QY 569 HisValPheGlnAlaLysValSerGluAspValAlaIleGlyThrLysValGlyAsnVal 588
 DB 1826 GAGAGTACAGAGCCAGGTGCTCCCATCAAGTCCCGCCAGCGGCTTCTGCTGACCATC 1885
 QY 589 ThrAlaLysAspProGluGlyLeuAspIleSerTyrrSerLeuArgLysAspGly 608
 DB 1886 CAGCCCTCCGACCCCACTCAAGCCGACACCTTCAAGTCTCTCAATGATCAAGAGGCC 1945
 QY 609 TyrLeuLysIleAspHisValThrGlyGluIlePheSerValAlaProLeuAsp---Arg 627
 DB 1946 TGGCTGTGATTAAGAAATTCCTCGGGGAGGTGCACACCGCCAGTCCCTGAGGGCGCC 2005
 QY 628 GlnAlaGlySerProTyrrArgValGlnValAlaIleThrGluValGlyLysSerSerLeu 647
 DB 2006 CACCTTGGGAGACCTTCAACAGGTGCTGTGAGAGCCAGAGTACAGCC----- 2053
 QY 648 SerSerValSerGluPheHisLeuIleLeuMetAspValAsnAspAsnProArgLeu 667
 DB 2054 -----CTGACTCTTGCCCTGTGCTGCCCTCCCA----- 2080
 QY 668 AlaLysAspTyrrThrGlyLeuPhePheCysHisProLeuSerAlaProGlySerLeuIle 687
 DB 2081 -----TACCTGTGACACCCCGGCAACACCATGCTGATTCGTG 2119
 QY 688 Phe---GlnAlaThrAspAspArgGlnHisLeuPheArgGlyProHisPheThrPheSer 706
 DB 2120 AGTGAAGCCAGCAAGAGCCCGATCTGGCCAGTGGGACAGCGTCC---TACACTTCAC 2176
 QY 707 LeuGly---SerGlySerLeuGlnAsnAspTyrGluValSerLysIleAsnGlyThrHis 725
 DB 2177 CTGTGTCCCAACCCAGGTGCAAGGATGCGCTTCAAGCTTCATGTGTTCCAT 2236
 QY 726 AlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaTyrrValValLeuIleArg 745

```

Db      2237 GCGTAACCTGACCTTGGCCCTGTCATTTGGGTGGAGCCACGTGAACATATATCCCGTGGTG 2296
Qy      746  ILeasnsapgllylgyArgProProleuGlulgllyleValserleuProvalThPheCys 765
Db      2297 GTCCAGCCCAATGCCCA-----ATGGGACAGCTCTGGTTCAGATGATGCTGTGT 2347
Qy      766  SerCys--ValglulgllySerCysPheArgProalaglYhiSglnThnglyleProthr 784
Db      2248 CGCTGCACAGTGGAGGGGCAATGCATGCGCAAGGTGGCCCGCATGAAGGCGCATGCCACG 2407
Qy      785  ValgluMetAlavalgllyleleuLeuThrThleuLeuValllleglylelleuAla 804
Db      2408 AAGCTGTGCGAGTGGGCATCTTGTAGGACCTGTGATGCAATAGGATCTTCCTCATC 2467
Qy      805  ValValPhelelAeArglllelylelyleAspIysgllylYsAspaenValgluSerAlaGlulAla 824
Db      2468 CTCATTTTCACCCACTGGACCATGTCAAGAAAGAACCCGATCAACAGCAGACAGC 2527
Qy      825  SergluVallys 828
Db      2528 GTGCCCTGGAAG 2539

RESULT 12
US-10-063-616-97
; Sequence 97, Application US/10063616
; Publication No. US20030013855A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,616
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 97
; LENGTH: 2848
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-616-97

Alignment Scores:
Pred. No.:      3,246-89      Length:      2848
Score:          908.50      Matches:      237
Percent Similarity: 50.51%      Conservative: 159
Best Local Similarity: 30.23%      Mismatches: 335
Query Match:    21.03%      Indels:      53
DB:              9          Gaps:      21

US-10-025-380-1081 (1-832) x US-10-063-616-97 (1-2848)
Qy      62  LeuThrglYglu-----ThraSpaenllePheVallllegluArgGlu---Gly 76
Db      296 CTGTGAGGAGGAGCAGGCAAGCAAGCAAGGCGCCATTGTGTATGATCCAGATTCGTGCG 355
Qy      77  LeuLeuTyrrYsnaArgAlaleuAspArgGluThrArgSerThrhisaenleuGlnVal 96
Db      356 TTCTGTGCTGTGACACAGAGGCGCTGTGACCGAGAGGAGCAGAGATCCAGCTACAGAGTC 415
Qy      97  AlAlaLeuAspAlaenGlyllelleValgluGllyProValProIleThrlleGluVal 116
Db      416 ACCCTGAGATGCCAGGATGATGCTTGTGGGGTCCACAGCTGTGCTTGTGCACGTG 475
Qy      117 LysAspIleAsnsapAenArgProThrPheleuGlnSerlySTyrglulyservalArg 136

```

```

Db      476 AAGATAGAGATGACCAAGTGGCCCATTTCTCCAGGCCATCTACAGAGCTCGGCTGAGC 535
Qy      137 GlnAsnserArgProgllylYsProPheLeuTyrrValAsnAlaThrAspLeuAspPro 156
Db      536 CGGGGTACCGAGGCTGGCATCCCTTCCTTCCTTAGGCTTCACAGCCGGGATGAGCCA 595
Qy      157  AlAthrProenGlylnleuTyrrYrglnlleValllleglnleuProMetIleAsnAsn 176
Db      596 GGCACAGCCCAATCTCGATCTTCGATTCACATCTCTGAGCCAGAGGCTCCAGCCACCTTCC 655
Qy      177  ValMetTyrrPheGlnlleAsnAsnlySThrGlyAlaAlaIleSerleuThraArgGlySer 196
Db      656 CAGACATGTTCCAGCTGAGGCGCTGGGCGCTGGCGCTGGCCCTCAGCCCAAGGGAGAGC 715
Qy      197  GlnGluLeuAsnProAlaYAsnProSerTyrrAsnleuValllleSerValYAspMet 216
Db      716 ACCAGCCTTACCAAGCCCTGAGAGGACCTACAGCTGTGGTACAGCTCAAGACATG 775
Qy      217  GlYglYlnSerGluAsnSerPheSerAspThrThSerValAspIlelleValThrglu 236
Db      776 GTGACCAAGGCC--TCAGGCCACACAGGCCACTGACCAAGTGAAGTCCATCATATAGC 832
Qy      237  AsnIleTyrrYsAlaProlylProValGluMetValGluAsnSerThraPheProhAsPro 256
Db      833 AGCAGCTGGGTGTCCTTAGAGCCTATCCAGCTGGAGAGATGCCAAAGTCCATATACCG 892
Qy      257  llelyleleThrglnValArgTyrrAsnAspProglYAlaGlnTyrrSerleuValAspLys 276
Db      893 CACCACATGGCCCAAGGTACCTGAGAGGGGTGATGTGACATATCACCTG----- 943
Qy      277  GlulYsleuProArgPheProPheSerleAspGlnGluGlyAspIleTyrrValThrgln 296
Db      944 GAGAGCCATCCCGGGGACCTTTGAAGTAAATGCAGAGGAAACCTTAGTACGACAGCA 1003
Qy      297  ProleuAspArgGlnGluYAspAlaTyrrValPheTyrrAlaValAlaYAspGluTyrr 316
Db      1004 GAGCTGACAGAGAGAGCCAGCGGTAGTACTCTGCTCCAGGTGGGGTCCAGATTCCTCAT 1063
Qy      317  GlyYsProleuSerTyrrProleuGluIleHisVallyleVallyAspIleleuAspAsn 336
Db      1064 GGGAGAGCATATGCGGCGCTCTGAGAGCTGACGTGCTGTGTAGTGAATGACCAAC 1123
Qy      337  ProProThrCysProSerProValThrhValPheGluValGlnGluAsnGluGly 356
Db      1124 GTGCTATCTGCTCCCTCCCGCTGACCCCAAGTACAGATCCCTAGCTCAGTCCACAGGT 1183
Qy      357  AsnserlleGlyThrhleuThralaHisAspArgAspGluGluAsnThralaAsnSerPhe 376
Db      1184 ACTGAAGTACATGACGTGCTGACAGAGATGACAGATGCCCGGCTCCCATTTCCAC 1243
Qy      377  LeuAsnTyrrArglleValgluGlnThrhProlyleuProMetAspGly--LeuPheleu 395
Db      1244 GTTGTGATCACTGCTGAGCCCTGAGCCCTGAGAGGAGATGAGGGGAGAGCCCTTCAG 1303
Qy      396  lleGlnThrhTyrrAlaGlyMetleuGlnleuAlaYlGlnSerleuYslys---GlnAsp 414
Db      1304 GTGAGCCCACTTCAGAGAGTGTGACGCTGGGGTCTCCCATCTCGAGAGGCCAGAAC 1363
Qy      415  ThrProGlnTyrrAsnleuThrhleGluValSerAspLysAsp-----PheTyrrleu 432
Db      1364 ATCCGCTTCTGTGCTGCTGCGCATGAGACCTGGCAGGGGCGAGAGGTGGCTTCACAGCAGC 1423
Qy      433  CysPheValGlnlleAsnVallleAspIleAsnAspGlnlleProIlePheGluYsSer 452
Db      1424 TGTGAATGCAATCCAGTACAGATATCAAGATCAAGCCCTGAGTTCATCATCTCC 1483
Qy      453  AspTyrrGlyAsnleuThrhleuAlaGluAspThrhAsnIleGlySerThrlleleuThrlle 472
Db      1484 CAGATTGGGCTATTAAGCTCTCCCTGAGAGTGTGAGCCCGGAGCTGTGGGCTATGCTA 1543
Qy      473  GlnAlaThrAspAlaAsp--GluProPheThrglySerSerlySleleu--TyrrHis 490

```

Db 1544 ACAAGCATGATGCTGACCTGAGCC-----GCTTCGCGCTCATGATTTTGCC 1594
 Qy 491 Tlellelysglyaspserygluylargleuylvalasphthasprohithasnthr 510
 Db 1595 ATTGAGGGGAGACAGAGAGGACTTTGGCTTGATGGAGCCA-----GACTCT 1648
 Qy 511 GltYrValIlellely-----LysProleuasphegluThraAlaIleValSerAsn 528
 Db 1649 GGGCATTTTACACTGACACTCTGCAAGAACCTCATTTATGAGAGAGCTCCAGTCAGAG 1708
 Qy 529 TleValPheIysalagluasnProgluProleuValPheglYValIysTyraAlaSer 548
 Db 1709 GTGGTGTGGTGTGGAGAGGTGGCGAGCTGGTG---GGGCGAGGCCAGGCCCTGGA 1765
 Qy 549 SerPheAlaIysPheThrIleuIleValThrasPValasgluAlaProgluInPheSerGln 568
 Db 1766 GCCACCCGACGAGTGTGCTGAGAGAGATGATGCCACCCCAAGTTGAGCAG 1825
 Qy 569 HisValPheGlnAlaIysValSerGluaspValAlaIleglYThrIysValIysAsnVal 588
 Db 1826 GAGAGTACAGAGGCGAGTGTCCCATCAGTGCAGCCCGCTCTTCTGCTGACATC 1885
 Qy 589 ThrAlaIysAspProgluIleuaspIleSerTyserIleuargIysAspThrargGly 608
 Db 1886 CAGCCCTCCGACCCCATCAGCCGAACCTCAGTTCTCCCTAGTCATGACTCAGAGGCG 1945
 Qy 609 TrpleuylIleasPheIysValIleThgluIlePheSerValAlaProleuAsp---Arg 627
 Db 1946 TGCGTCTGCAATGGAAATTCCTCGGAGGTGACACCGCCAGCTCTGACAGGGCGCC 2005
 Qy 628 GluAlaIysSerProIyrrargValIgluValAlaIleThrgluValIglYIysSerIleu 647
 Db 2006 CAGCTGGGACACTTACAGGCTCTGTGAGGAGCCAGATACAGCC----- 2053
 Qy 648 SerSerValSerIupheIleuIleuWeraspValasnaPasnProIargIleu 667
 Db 2054 -----CTGACTCTGCGCTGCGCTGCCCA----- 2080
 Qy 668 AlaIysAspTyrrThrglyLeuPhePheCysIleProleuSerAlaProglYSerIleuIle 687
 Db 2081 -----TACCTCGACACCCCGCCCAAGACATGCTGATGTG 2119
 Qy 688 Phe---GluAlaThrAspAspAspGlnIleuPheargGlyProIAspPheSer 706
 Db 2120 AGTGAAGCCAGACAGACCCGATCTGGCCAGTGGACAGCTCC---TACAGCTTACCC 2176
 Qy 707 LeuGlyI---SerGlySerIleuGlnAsnAspTrpGluValSerIysIleasnglyThris 725
 Db 2177 CTGTGTCCCAACCCCAACGCGTCAACGGGATGGCGCTCCAGACTCTCAATGATGCCAT 2236
 Qy 726 AlaIargIleuSerThrIargIleThrasPhegluIleuIargIleValIleuIlearg 745
 Db 2237 GCGTAACCTCACCTGGCCCTGATGGGTGAGCGACATGAACATATATCCCGTGGTG 2296
 Qy 746 IleasnAspGlyIargProIleuGlnIleValSerIleProValIleThrasPheCys 765
 Db 2297 GTGAGCCACATGCCAG-----ATGAGCAGCTCTCGGTTCAGATGATGTGTGT 2347
 Qy 766 SerCys---ValGluIysSerCysPheargProIaIleGlnIleIleuIleProthr 784
 Db 2348 CGTGAACAGTGGAGGGGAGTGCATGCGCAAGTGGGCCGCAATGAAGGGCATGCCACG 2407
 Qy 785 ValIleuMetAlaValIleIleuIleuThrIleuIleuValIleGlyIleIleuAla 804
 Db 2408 AAGCTGCGCAGTGGGATCTTGTAGCACCTGTGAGCAATGAGCAATTTCTCATTC 2467
 Qy 805 ValIleuPheIleargIleIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 824
 Db 2468 CTCATTTTACCCACTGAGCATGTCAAGAGAGAGACCCGATCAACAGCAGACAGC 2527
 Qy 825 SerGluValIys 828
 Db 2528 GTGCCCTCGAAG 2539

RESULT 13
 US-10-063-502-97
 ; Sequence 97, Application US/10063502
 ; Publication No. US20030023042A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gerltsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Matanabe, Colin K.
 ; APPLICANT: Wood, William L.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3230R1C1
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; CURRENT APPLICATION NUMBER: US/10/063,502
 ; CURRENT FILING DATE: 2002-05-01
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 170
 ; SEQ ID NO 97
 ; LENGTH: 2848
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-063-502-97
 Alignment Scores:
 Pred. No.: 3,246-89 Length: 2848
 Score: 908.50 Matches: 237
 Percent Similarity: 50.51% Conservative: 159
 Best Local Similarity: 30.23% Mismatches: 335
 Query Match: 21.03% Indels: 53
 DB: 9 Gaps: 21
 US-10-025-380-1081 (1-832) x US-10-063-502-97 (1-2848)
 Qy 62 LeuThrglyGlu-----ThrasPasnIlePheValIleGluThrgluIleu---Gly 76
 Db 296 CTGTCAAGGAGCTACAGGCAAGCAAGTGCAGGCGCCATTTGCTATGATGATCCAGATTTGCG 335
 Qy 77 LeuIleuTyrrAsnArgAlaIleuAspArgGluThrasPheSerThrisAsnIleuIleVal 96
 Db 356 TTCTGTGTGTGACACAGGCGCTCGACCGAGAGACAGACAGATACAGATACAGATGTC 415
 Qy 97 AlaIleuAspAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 116
 Db 416 ACCCTGAGATGACAGATGACATGCTGTGTGGGTCCACAGCTGTGTGTGACGTG 475
 Qy 117 LysAspIleAsnAspAsnAspProthrPheIleuGlnSerIysTyrrGluIysSerValArg 136
 Db 476 AAGATGAGATACACAGAGTGCCTTTCTCAAGCCATTTACAGAGCTCGGCTGAGC 535
 Qy 137 GlnAsnSerArgProglYIysProPheIleuTyrrValAsnAlaThrasPleuAspPro 156
 Db 536 CGGGGTACACAGGCTGGCATCCCTTCTCTGAGGCTTCAAGCCGAGTGAAGCA 595
 Qy 157 AlaThrasPheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 176
 Db 596 GGCACAGCCAACTCGGATCTTCCATTCACATCTGAGCCAGGCTCCAGCCAGCCCTTC 655
 Qy 177 ValMetTyrrPheGlnIleuAsnAsnIleuIleuIleuIleuIleuIleuIleuIleuIleu 196
 Db 656 CAGACATGTTCCAGCTGAGCTCGGCTGGGGGCTTGCCCTTCAGCCCAAGGGAGC 715
 Qy 197 GlnIleuIleuAsnProIalysAsnProSerTyrrAsnIleuValIleuSerValIysAspMet 216
 Db 716 ACCAGCTTACACAGCCCTCGAGAGAGACCTTACAGCTGTGTGTACAGGTCAAGGACATG 775
 Qy 217 GlyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 236
 Db 776 GTGACCAAGGCC---TCAGGCCACAGGCCACATGCAACGCTGAGAGTTCATCATATAGAG 832

OY	237	Asn1IetPpLyAlAProLyserProValGluMetValGluAsnSerThrAspProHisPro	256
Db	833	AGCACTGGGTGTCCTCCAGACTTCCACCTCGCAGAGAAATCTCAAAAGTCTTATCCCG	892
OY	257	IleLyS1IethrGlnValArgTrpAsnAspProGlyAlaGlnTrpSerLeuValAspLys	276
Db	893	CACCAACATGGCCCAAGTACACTGAGATGGGGGTGATGTGCATTAACACTG-----	943
OY	277	G1uLyS1IeuProArgPheProPheSerIleAspGlnGluYAspIleTyrValThrGln	296
Db	944	GAGAGCCATCCCCCGGAGCCCTTTAAGTAATGCAAGAGGAAACCTTACGTGCACAGA	1003
OY	297	ProLeuAspArgGluGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyr	316
Db	1004	GAGCTGGACAGAGAAAGCCAGCGTGAATACCTGCTCCAGAGTGGCGGCTCAGAAATTCAT	1065
OY	317	GlyLysProLeuSerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsn	336
Db	1064	GGCGGAGACTATGCGCGCCCTCTTGAGGTGCACGTGCTGTGATGAAGAGAAATGACAAC	1123
OY	337	ProProThrCysProSerProValThrValPheGluValGlnGluAsnGluArgLeuGly	356
Db	1124	GTGCTTATCTGCTCTCCCGTGAACCCACAGTCAGACATTCCTGAAGTCAATCCACAGGT	1183
OY	357	AsnSerIleGlyThrLeuThrAlaHisAspArgAspGluAsnThrAlaAsnSerPhe	376
Db	1184	ACTGAAGTACTAGACTGTCTGACAGCAAGAGATGCAGATGCCCGCGCTCCCAATTCCAC	1243
OY	377	LeuAsnTyrArgIleValGluGlnThrProLysLeuProMetAspGly--LeuPheLeu	395
Db	1244	GTGTGTATCAGCTCTTCAAGCCCTGAGCTTGAGATGGGGTGAAGAGGGAGAGCTTCCAG	1303
OY	396	IleGlnThrTyrAlaGluMetLeuGlnLeuAlaLysGlnSerLeuLysLys--GlnAsp	414
Db	1304	GTGAGACCCCATCTCAGCAGAGTGAACGTGGGGGTGCTCCCATCTCCAGCAGGCCAGAAC	1363
OY	415	ThrProGlnTyrAsnLeuThrIleGluValSerAspLysAsp----PheLysThrLeu	432
Db	1364	ATCTGCTTCTGTGCTGCTGACATGAGCCTGGCAGCGCAGAGAGGTGCTTCAGCAGCAGC	1423
OY	433	CysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLysSer	452
Db	1424	TGTGAAGTCGAAGTCCGACATCAGATATCAATGATACAGCCCTGAGTTATATCATCTCC	1483
OY	453	AspTyrGlyAsnLeuThrLeuAlaGluAspThrAsnIleGlySerThrIleLeuThrIle	472
Db	1484	CAGATTGGGCTTATAGGCTCCCTCGTAGATGTGAGCCCGGAGATCTGTGGCATGTGTA	1543
OY	473	GlnAlaThrAspAlaAsp--GluProPheThrGlySerSerLysIleLeu--TyrHis	490
Db	1544	ACAGCGATTGATGCTGACCTCGAGCC-----GCCCTTCGGCTCATGAAATTTGCC	1594
OY	491	IleIleLysGlyAspSerGluGlyArgLeuGlyValAlaAspThrAspProHisIleThrAsnThr	510
Db	1595	ATTGAGAGGGGAGACACAGAAAGGACTTTGGCTCGATTTGGAGCCA-----GACTCT	1648
OY	511	GlyTyrValIleIleLys----LysProLeuAspPheGluThrAlaAlaValSerAsn	528
Db	1649	GGGCAATTGAACCTCAGACTCTGCAGAACTCATATTAGAGGACGCTCCAAAGTCATGAG	1708
OY	529	IleValPheLysAlaGluAsnProGluProLeuValPheGlyValLysTyrAsnAlaSer	548
Db	1709	GTGTTGTGGTGTGTCAGAGTGTGGCCAGCTGTGGT--GGGCCAGGCCCAAGGCTTGGGA	1765
OY	549	SerPheAlaLysPheThrLeuIleValThrAspValAsnGluAlaProGlnPheSerGln	568
Db	1766	GCCACCGGCAAGGTACTGTCTGATGGAGAGAGTGAATGCACACCCCAAGTTGGACAG	1825
OY	569	HisValPheGlnAlaLysValSerGluAspValAlaIleGlyThrLysValGlyAsnVal	588
Db	1826	GAGAGCTACGAGGCGCAAGTGTCCCATCATGCGCCCAAGCCGGCTCTTCTCTGTGACATCT	1885

```

QY 589 thralAlayAspProGluGlyLeuAspLLeSerTySerLeuAlaGlyAspThrArgGly 608
Db 1886 CAGCCCTCCGACCCCATGATGACCCGAAACCTCCAGGTTCTCCCTTATGATGACTCGAGAGGC 1945
QY 609 TrpleuLysIleAspHisValThrGlyGluIlePheSerValAlaProLeuAsp---Arg 627
Db 1946 TGGCTCTGCATGAGAAATTCCTCGGGGAGGTGACACCGCCAGTCCCTCGAAGGCGCC 2005
QY 628 GluAlaGlySerProTyTArgValGlnValAlaThrGluValGlyIlySerSerIeu 647
Db 2006 CAGCTCGGGGACACCTCACCGGTCTGTGGAGGCCCGCAGATACAGCC----- 2053
QY 648 SerSerValSerGluPheHisIleuIleLeuMetAsuValaLanAspAspProPArgIeu 667
Db 2054 -----CTAATCTTTGCCCCCTGTGCCCCCTCCAA----- 2088
QY 668 AlaLysAspTyTrhIhGlyLeuPhePheCysHisProLeuSerAlaProGlySerIeuIle 687
Db 2081 -----TACCTCTGCACACCCCGCCCAAGACACATGGCTTATGCTG 2119
QY 688 Phe---GluAlaThrAspAspAspGlnHisIleuPheArgGlyProHisPheThrPheSer 706
Db 2120 AGTGAGACCCGACAGACAGCCCGATCTGGCCAGTGAGGACCGGTCC---TACAGCTTCACC 2176
QY 707 LeuGly---SerGlySerLeuGlnAspTrpGluValSerLysIleAsnGlyThrHis 725
Db 2177 CTGTGTCCCAACCCCGACGTGCACACGGGATTTGGCGCTCCAGACTCTCAATGGTTCCAT 2236
QY 726 AlaArgLeuSerThrThrArgHisIleThrAspPheGluGluArgAlaTyValValIleuIleArg 745
Db 2237 GCCTAACCTCACCTTGGCCCTGATGGGTGAGGACGACCATGAACATATATCCCGGTGTG 2296
QY 746 IleAsnAspGlyArgProProLeuGluGluIlyIleValSerLeuProValThrPheCys 765
Db 2297 GTGAGCCACATGCGCCAG-----ATGGGACAGCTCGTTCAGTGATGCTGTGT 2347
QY 766 SerCys---ValGluGlySerCysPheArgProAlaGlyHisGlnThrGlyIleProThr 784
Db 2348 CGGTGCACACGTGAGGGGACAGTGTGATCGGCAAGGTGGGCGCATGAAGGGCATGCCACG 2407
QY 785 ValGlyMetAlaValGlyIleLeuLeuThrThrIeuPheValIleGlyIleIleLeuAla 804
Db 2408 AAGCTGTCGGCAGTGGGATCTTTGTGGACCTCTGGTAGCAATAGGAATTTCTCTATC 2467
QY 805 ValValPheIleArgIleLysLysAspLysGlyLysAspAsnValGluSerAlaGlnAla 824
Db 2468 CTCATTTTCAACCCACGTCGACCATGTCAAGAGAAGAGACCCGATCAACGACGACAGCG 2527
QY 825 SerGluValLys 828
Db 2528 GTGCCCTCGAAG 2539

RESULT 14
US-10-227-884-187
; Sequence 187, Application US/10227884
; Publication No. US20030027988A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Auctin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P350P1C79
; CURRENT APPLICATION NUMBER: US/10/227,884
; CURRENT FILING DATE: 2002-08-26

```


1	PRIOR APPLICATION NUMBER: 10/119,480
2	PRIOR FILING DATE: 2002-04-09
3	PRIOR APPLICATION NUMBER: 60/059113
4	PRIOR FILING DATE: 1997-09-17
5	PRIOR APPLICATION NUMBER: 60/062287
6	PRIOR FILING DATE: 1997-10-17
7	PRIOR APPLICATION NUMBER: 60/063549
8	PRIOR FILING DATE: 1997-10-28
9	PRIOR APPLICATION NUMBER: 60/064103
10	PRIOR FILING DATE: 1997-10-31
11	PRIOR APPLICATION NUMBER: 60/065873
12	PRIOR FILING DATE: 1997-12-17
13	PRIOR APPLICATION NUMBER: 60/078910
14	PRIOR FILING DATE: 1998-03-20
15	PRIOR APPLICATION NUMBER: 60/079294
16	PRIOR FILING DATE: 1998-03-25
17	PRIOR APPLICATION NUMBER: 60/079656
18	PRIOR FILING DATE: 1998-03-26
19	PRIOR APPLICATION NUMBER: 60/079728
20	PRIOR FILING DATE: 1998-03-27
21	PRIOR APPLICATION NUMBER: 60/081819
22	PRIOR FILING DATE: 1998-04-15
23	PRIOR APPLICATION NUMBER: 60/081955
24	PRIOR FILING DATE: 1998-04-15
25	PRIOR APPLICATION NUMBER: 60/082804
26	PRIOR FILING DATE: 1998-04-22
27	PRIOR APPLICATION NUMBER: 60/084441
28	PRIOR FILING DATE: 1998-05-06
29	PRIOR APPLICATION NUMBER: 60/085323
30	PRIOR FILING DATE: 1998-05-13
31	PRIOR APPLICATION NUMBER: 60/085579
32	PRIOR FILING DATE: 1998-05-15
33	PRIOR APPLICATION NUMBER: 60/086392
34	PRIOR FILING DATE: 1998-05-22
35	PRIOR APPLICATION NUMBER: 60/089532
36	PRIOR FILING DATE: 1998-06-17
37	PRIOR APPLICATION NUMBER: 60/089538
38	PRIOR FILING DATE: 1998-06-17
39	PRIOR APPLICATION NUMBER: 60/089905
40	PRIOR FILING DATE: 1998-06-18
41	PRIOR APPLICATION NUMBER: 60/090472
42	PRIOR FILING DATE: 1998-06-24
43	PRIOR APPLICATION NUMBER: 60/090557
44	PRIOR FILING DATE: 1998-06-24
45	PRIOR APPLICATION NUMBER: 60/090691
46	PRIOR FILING DATE: 1998-06-25
47	PRIOR APPLICATION NUMBER: 60/090695
48	PRIOR FILING DATE: 1998-06-25
49	PRIOR APPLICATION NUMBER: 60/091982
50	PRIOR FILING DATE: 1998-07-07
51	PRIOR APPLICATION NUMBER: 60/095302
52	PRIOR FILING DATE: 1998-08-04
53	PRIOR APPLICATION NUMBER: 60/095318
54	PRIOR FILING DATE: 1998-08-04
55	PRIOR APPLICATION NUMBER: 60/095916
56	PRIOR FILING DATE: 1998-08-10
57	PRIOR APPLICATION NUMBER: 60/096146
58	PRIOR FILING DATE: 1998-08-11
59	PRIOR APPLICATION NUMBER: 60/096791
60	PRIOR FILING DATE: 1998-08-17
61	PRIOR APPLICATION NUMBER: 60/097986
62	PRIOR FILING DATE: 1998-08-26
63	PRIOR APPLICATION NUMBER: 60/098544
64	PRIOR FILING DATE: 1998-08-31
65	PRIOR APPLICATION NUMBER: 60/099596
66	PRIOR FILING DATE: 1998-09-09
67	PRIOR APPLICATION NUMBER: 60/099598
68	PRIOR FILING DATE: 1998-09-09
69	PRIOR APPLICATION NUMBER: 60/099803
70	PRIOR FILING DATE: 1998-09-10
71	PRIOR APPLICATION NUMBER: 60/099811
72	PRIOR FILING DATE: 1998-09-10
73	PRIOR APPLICATION NUMBER: 60/099812

PRIOR FILING DATE:	1998-09-10
PRIOR APPLICATION NUMBER:	60/0399816
PRIOR FILING DATE:	1998-09-10
PRIOR APPLICATION NUMBER:	60/100038
PRIOR FILING DATE:	1998-09-11
PRIOR APPLICATION NUMBER:	60/100385
PRIOR FILING DATE:	1998-09-15
PRIOR APPLICATION NUMBER:	60/100390
PRIOR FILING DATE:	1998-09-15
PRIOR APPLICATION NUMBER:	60/100627
PRIOR FILING DATE:	1998-09-16
PRIOR APPLICATION NUMBER:	60/100848
PRIOR FILING DATE:	1998-09-18
PRIOR APPLICATION NUMBER:	60/100919
PRIOR FILING DATE:	1998-09-17
PRIOR APPLICATION NUMBER:	60/101477
PRIOR FILING DATE:	1998-09-23
PRIOR APPLICATION NUMBER:	60/101738
PRIOR FILING DATE:	1998-09-24
PRIOR APPLICATION NUMBER:	60/101741
PRIOR FILING DATE:	1998-09-24
PRIOR APPLICATION NUMBER:	60/101916
PRIOR FILING DATE:	1998-09-25
PRIOR APPLICATION NUMBER:	60/101916
PRIOR FILING DATE:	1998-09-24
PRIOR APPLICATION NUMBER:	60/101922
PRIOR FILING DATE:	1998-09-24
PRIOR APPLICATION NUMBER:	60/106178
PRIOR FILING DATE:	1998-10-28
PRIOR APPLICATION NUMBER:	60/106248
PRIOR FILING DATE:	1998-10-29
PRIOR APPLICATION NUMBER:	60/106454
PRIOR FILING DATE:	1998-10-30
PRIOR APPLICATION NUMBER:	60/106905
PRIOR FILING DATE:	1998-11-03
PRIOR APPLICATION NUMBER:	60/108787
PRIOR FILING DATE:	1998-11-17
PRIOR APPLICATION NUMBER:	60/108801
PRIOR FILING DATE:	1998-11-17
PRIOR APPLICATION NUMBER:	60/108849
PRIOR FILING DATE:	1998-11-18
PRIOR APPLICATION NUMBER:	60/112422
PRIOR FILING DATE:	1998-12-15
PRIOR APPLICATION NUMBER:	60/113266
PRIOR FILING DATE:	1998-12-22
PRIOR APPLICATION NUMBER:	60/113605
PRIOR FILING DATE:	1998-12-23
PRIOR APPLICATION NUMBER:	60/113621
PRIOR FILING DATE:	1998-12-23
PRIOR APPLICATION NUMBER:	60/115568
PRIOR FILING DATE:	1999-01-12
PRIOR APPLICATION NUMBER:	60/115565
PRIOR FILING DATE:	1999-01-12
PRIOR APPLICATION NUMBER:	60/123618
PRIOR FILING DATE:	1999-03-10
PRIOR APPLICATION NUMBER:	60/125259
PRIOR FILING DATE:	1999-03-19
PRIOR APPLICATION NUMBER:	60/125775
PRIOR FILING DATE:	1999-03-23
PRIOR APPLICATION NUMBER:	60/126773
PRIOR FILING DATE:	1999-03-29
PRIOR APPLICATION NUMBER:	60/127887
PRIOR FILING DATE:	1999-04-05
PRIOR APPLICATION NUMBER:	60/130232
PRIOR FILING DATE:	1999-04-21
PRIOR APPLICATION NUMBER:	60/131022
PRIOR FILING DATE:	1999-04-26
PRIOR APPLICATION NUMBER:	60/131270
PRIOR FILING DATE:	1999-04-27


```

PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100038
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101786
PRIOR FILING DATE: 1998-09-25
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106178
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/106248
PRIOR FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: 60/106464
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 60/106905
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/108787
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108801
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108849
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/125259
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 60/125775
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/126773
PRIOR FILING DATE: 1999-03-29

PRIOR APPLICATION NUMBER: 60/127887
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/130232
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131022
PRIOR FILING DATE: 1999-04-26
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131445
PRIOR FILING DATE: 1999-04-28
PRIOR APPLICATION NUMBER: 60/134287
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/140650
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/140723
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146963
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/149320
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149638
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/151733
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/164418
PRIOR FILING DATE: 1999-11-09
PRIOR APPLICATION NUMBER: 60/166361
PRIOR FILING DATE: 1999-11-16
PRIOR APPLICATION NUMBER: 60/169445
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169835

Alignment Scores:
Pred. No.: 3,24e-89 Length: 2848
Score: 908.50 Matches: 237
Percent Similarity: 50.51% Conservative: 159
Best Local Similarity: 30.23% Mismatches: 335
Query Match: 21.03% Indels: 53
DB: 9 Gaps: 21

US-10-025-380-1081 (1-832) x US-10-230-163-187 (1-2848)
QY 62 LeuThrglyGlu-----PhraSpaaniIephValIIeGluArgiu---Gly 76
DB 296 CTCTCGAGGAGCTCAGGCAAGCACTGAGGCCCATTTGTAATGATCCAGATTTCGCC 355
QY 77 LeuLeuTyTyraSnaIrgAlaIeuaSpaIrgIurIarSeraThrhiaSnaIeuaI 96
DB 356 TTCCTGCTGTGACCGAGGCGCTTGAGCGAGAGGACGAGACGATCAGCTACAGTGC 415
QY 97 AlaAlaIeuaSpAlaasnglyIleIleValIguIyProValProIleThrlIeGluVal 116
DB 416 ACCCTGAGATGACAGATGACATGCTGTGGGTCCACACCTGTGCTTGACACGTC 475
QY 117 LysAspIleAsnaSpaSnaIrgProThrPheIeugInserLyTyTgIuIySeraValArg 136
DB 476 AAGGATGAGAAATGACCAAGGTGCCCATTTCTCAAGCATCTTACAGAGCTCGCGTAC 535
QY 137 GlinaSeraIrgProGlyLysProPheLeuTyTyraIaSnaIaThraSnaIeuaSpaPro 156
DB 536 CGGGGTACCGAGCTGAGCATCCCTTCCTTCTTGAGGCTTCAAGCCGGGATGAGCCA 595
```

QY 157 AlaThrProAsnGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsn 176
 DB 596 GGCACACCACTGGATCTTCGATCCACATCTGAGCCAGGCTCCAGCCCACTTC 655
 QY 177 ValMetTyrPheGlnIleAsnAsnLeuThrGlyAlaIleSerLeuThrArgGlnGlySer 196
 DB 656 CCAGACATGTTCCAGCTGAGACCTCGGCTGGGGCTCTGGCCCTCCACCCCAAGGGAGC 715
 QY 197 GlnGlnLeuAsnProAlaValAsnProSerTyrAsnLeuValIleSerValIleAspMet 216
 DB 716 ACCAGCTTGACCAAGCTGAGAGACCTACAGCTGTGGTACAGGTCAAGAGACATG 775
 QY 217 GlyGlnIleSerGlnAsnSerPheSerAspThrThrSerValAspIleIleValThrGln 236
 DB 776 GGTACACAGAGCC--TCAGGCCACAGGCCACCTCCACCTGGAGAGTCCATCATAGAG 832
 QY 237 AsnIleTrpValAspProValGlnMetValGlnAsnSerThrAspProHisPro 256
 DB 833 AGCAGCTGGGTGCTCCCTAGACCTATCCAGCTGGAGAGATCCAAAGTCTTACCCG 892
 QY 257 IleValIleThrGlnValArgTrpAsnAspProGlyAlaGlnTyrSerLeuValAspLys 276
 DB 893 CACACATGGCCCAAGTACACTGGAGGTGATGTGCACTATCACCTG-- 943
 QY 277 GlnLysLeuProArgPheProPheSerIleAspGlnGlnGlyAspIleTyrValThrGln 296
 DB 944 GAGGACCATCCCGGGGACCTTGAAGTGAATGACAGAGGGAACCTTCAGTACAGAGA 1003
 QY 297 ProLeuAspArgGlnGlnLysAspAlaTyrValPheTyrAlaValAlaLysAspGlnTyr 316
 DB 1004 GAGCTGGACAGAGAACCCAGGCTGAGTACTGCTCCAGAGTGGCGGCTCAGAAATTCCT 1063
 QY 317 GlyLysProLeuSerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsn 336
 DB 1064 GCGAGAGCATATGGCGCCCTCTGAGCTGACAGCTGCGTGAATGAGATGAGAAATGACAA 1123
 QY 337 ProProThrCysProSerProValThrValPheGlnValGlnGlnAsnGlnArgLysGln 356
 DB 1124 GTGCTTATCTGCTCCCTCCGCTGACCCCACTGACAGCTCCCTGAGCTCAGTCCACAGGT 1183
 QY 357 AsnSerIleGlyThrLeuThrAlaHisAspArgAspGlnGlnAsnThrAlaAsnSerPhe 376
 DB 1184 ACTGAAGTACTAGACTCTCAGCAGAGAGATGACAGTCCCGGCTCCCAATTCCTCAC 1243
 QY 377 LeuAsnTyrArgIleValGlnGlnThrProLysLeuProMetAspGly--LeuPheLeu 395
 DB 1244 GTTGTGATACAGCTCTGAGCCCTGAGCCTGAGATGGGATGAGGGGAGAGAGCTTCAG 1303
 QY 396 IleGlnIleThrAlaGlnMetLeuGlnIleAlaLysGlnSerLeuLysLys--GlnAsp 414
 DB 1304 GTGACCCCACTTACAGGAGTGTGACGCTGGGGGTCTCCACCTCCAGCAGGCGCAGAAC 1363
 QY 415 ThrProGlnTyrAsnLeuThrIleGlnValSerAspLysAsp--PheLysThrLeu 432
 DB 1364 ATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423
 QY 433 CysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGlnLysSer 452
 DB 1424 TGTGAAGTCGAAGTCGAGTACAGATATCATATGATCAGCCCTGAGTTCATCATCTCC 1483
 QY 453 AspTyrGlnLeuLeuThrLeuAlaGlnAspThrAsnIleGlySerThrIleLeuThrIle 472
 DB 1484 CAGATGGGCTTAAAGCTCTCTGAGGATGTGAGCCCGGAGCTGCTGCTGCTGCTGCTGCT 1543
 QY 473 GlnAlaThrAspAlaAsp--GlnProPheThrGlySerSerLysIleLeu--TyrHis 490
 DB 1544 ACAAGCCATGATGCTGAGCTGAGCC--GCTTCCGCTCATGATGATTTTCC 1594
 QY 491 IleIleLysGlnAspSerGlnGlyArgGlnGlnValAspThrAspProHisThrAsnThr 510
 DB 1595 ATTGAGAGGGAGACAGAAAGGAGCTTTGGCTGATGTGGAGCCA--GACTCT 1648

QY 511 GlyTyrValIleIleLys-----LysProLeuAspPheGlnThrAlaValSerAsn 528
 DB 1649 GGCATCTTATGACTGACACTGACAGAGCTGAGTATGAGGACAGCTCCAGTATGAG 1708
 QY 529 IleValPheLysAlaGlnAsnProGlnProLeuValPheGlyValLysTyrAsnAlaSer 548
 DB 1709 GTGGTGTGATGTGATGAGTGTGAGGAGCTGTGT--GGGCGAGCCCAAGCCCTGGA 1765
 QY 549 SerPheAlaLysPheThrLeuIleValThrAspValAsnGlnAlaProGlnPheSerGln 568
 DB 1766 GCCACCGCCAGCGGTGCTGTGTGAGTGAAGAGTATGACACCCCAAGTTCAGCCAG 1825
 QY 569 HisValPheGlnAlaLysValSerGlnAspValAlaIleGlyThrLysValGlnAsnVal 588
 DB 1826 GAGAGCTAGAGGACAGTGTCCCATCAGTGGCCCAAGCGGCTCTTCTGCTGAGCATT 1885
 QY 589 ThrAlaLysAspProGlnGlnLysAspIleSerTyrSerLeuArgLysThrArgGly 608
 DB 1886 CAGCCCTCGACCCCTTCAGCCGAGACCTCAGGTTCTCTGATCAATGATCAGAGGC 1945
 QY 609 TrpLeuLysIleAspHisValThrGlnGlnIlePheSerValAlaProLeuAsp--Arg 627
 DB 1946 TGGCTGTGATGAGAAATTCCTCGGAGAGTACACAGCCCACTGCTGAGGGCGCC 2005
 QY 628 GlnAlaGlySerProTyrArgValGlnValAlaIleThrGlnValGlyLysSerLeu 647
 DB 2006 CAGCTGGGAGACCTACACAGTGTGTGAGAGCCAGATTCAGCC----- 2053
 QY 648 SerSerValSerGlnPheHisLeuIleMetAspValAsnAspAsnProAlaGln 667
 DB 2054 -----CTGACTCTGCTCCCTGCTGCTCCCA----- 2080
 QY 668 AlaLysAspTyrThrGlnLeuPhePheCysHisProLeuSerAlaProGlySerLeuIle 687
 DB 2081 -----TACTCTGACACCCCGCCAGACCATGCTGATCTG 2119
 QY 688 Phe--GlnAlaThrAspAspArgGlnIleLeuPheArgGlyProHisPheThrPheSer 706
 DB 2120 AGTGAACCAAGCAAGAGCCCGATCGGCGAGTGGGAGGAGGTCC--TACAGCTTCAC 2176
 QY 707 LeuGly--SerGlySerLeuGlnAsnAspTrpGlnValSerLysIleAsnGlnThrHis 725
 DB 2177 CTGTGTCCCAACCCAGCGTGAACGAGTGTGCGGCTCCAGACTCTCATGTGTTCCAT 2236
 QY 726 AlaArgLeuSerThrArgHisThrAspPheGlnGlnAlaTyrValValIleuLeuArg 745
 DB 2237 GCTTACTTACCTTGGCCCTGCACTGGGTGAGCCACGTGAACATATCCCGGTG 2296
 QY 746 IleAsnAspGlyLysArgProProLeuGlnGlyIleValSerLeuProValThrPheCys 765
 DB 2297 GTACAGCCAAATGCCAG-----ATGTGCACTCTGTTGAGATGATCGTGT 2347
 QY 766 SerCys--ValGlnGlySerCysPheArgProAlaGlnHisGlnThrGlyIleProThr 784
 DB 2348 CGCTGCAACCTGAGGGGCGAGTCAATCCGAGGTGGCGCGCATGAAGGACAGCCAGC 2407
 QY 785 ValGlyMetAlaValGlyIleLeuLeuThrThrLeuValIleGlyIleIleValAla 804
 DB 2408 AACCTGTGGCAATGGGCACTTGTGAGCACTGTGACCAATGAGATCTTCTCATC 2467
 QY 805 ValValPheIleArgIleLysLysAspLysGlyLysAspAsnValGlnSerAlaGlnAla 824
 DB 2468 CTGATTTTCAACCACTGAGACATGTCAAGGAAGAAGACCGGATCAACAGAGACAGC 2527
 QY 825 SerGlnValLys 828
 DB 2528 GTGCGCTGAAG 2539

Search completed: June 22, 2003, 03:29:38
 Job time : 417 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2003, 01:09:40 ; Search time 114 Seconds
(without alignments)
2238.202 Million cell updates/sec

Title: US-10-025-380-1081

Perfect score: 4321
Sequence: 1 MIQAHLSLCLMLYATG.....DKGKNVSAQSEVPLRS 832

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xid
-O=/cgnt2_1/USPTO_spool/US10025380/rnat_17062003_173328_24692/app_query.fasta_1.1031
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MNTLEN=0 -MAXLEN=200000000
-USER=US10025380 @CGN 1.1 39 @rnat 17062003_173328_24692 -NCPUI=6 -ICPI=3
-NO MAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -KGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

- 1: /cgnt2_6/ptodata/2/ina/5A COMB.seq:*
- 2: /cgnt2_6/ptodata/2/ina/5B COMB.seq:*
- 3: /cgnt2_6/ptodata/2/ina/6A COMB.seq:*
- 4: /cgnt2_6/ptodata/2/ina/6B COMB.seq:*
- 5: /cgnt2_6/ptodata/2/ina/PCTUS COMB.seq:*
- 6: /cgnt2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4321	100.0	2499	1	US-08-431-560-2
2	4321	100.0	2499	1	US-08-463-345-2
3	684.5	15.8	3048	1	US-08-188-228-47
4	684.5	15.8	3048	1	US-08-332-638-41
5	684.5	15.8	3048	1	US-08-332-638-47
6	671.5	15.5	3552	4	US-09-643-597-126
7	660.5	15.3	2690	1	US-08-188-228-61
8	660.5	15.3	2690	1	US-08-332-643-55
9	660.5	15.3	2690	1	US-08-332-638-61
10	630.5	14.6	2779	1	US-08-474-067-3
11	630.5	14.6	2779	2	US-08-474-068A-3
12	630.5	14.6	2779	2	US-08-472-481-3

13	630.5	14.6	3959	1	US-08-474-067-1	Sequence 1, Appl1
14	630.5	14.6	3959	2	US-08-474-068A-1	Sequence 1, Appl1
15	630.5	14.6	3959	2	US-08-472-481-1	Sequence 1, Appl1
16	593	13.7	2808	1	US-08-237-919-1	Sequence 1, Appl1
17	593	13.7	2808	4	US-08-732-429-1	Sequence 1, Appl1
18	593	13.7	2808	4	US-09-798-267-1	Sequence 1, Appl1
19	593	13.7	2808	5	PCT-US95-05518-1	Sequence 1, Appl1
20	574.5	13.3	3581	2	US-08-738-349-1	Sequence 1, Appl1
21	563	13.0	346	4	US-09-221-298-21	Sequence 21, Appl1
22	562.5	13.0	3712	4	US-08-738-349-3	Sequence 3, Appl1
23	560	13.0	627	4	US-09-385-982-514	Sequence 514, App
24	555.5	12.9	2625	1	US-08-188-228-57	Sequence 57, Appl1
25	555.5	12.9	2625	1	US-08-332-643-51	Sequence 51, Appl1
26	555.5	12.9	2625	1	US-08-332-638-57	Sequence 57, Appl1
27	550.5	12.7	3136	1	US-08-188-228-41	Sequence 41, Appl1
28	550.5	12.7	3136	1	US-08-332-638-41	Sequence 41, Appl1
29	546.5	12.6	2550	1	US-08-188-228-53	Sequence 53, Appl1
30	546.5	12.6	2550	1	US-08-332-643-47	Sequence 47, Appl1
31	546.5	12.6	2550	1	US-08-332-638-53	Sequence 53, Appl1
32	545.5	12.6	3914	2	US-08-738-349-5	Sequence 5, Appl1
33	532	12.3	2521	1	US-08-188-228-59	Sequence 59, Appl1
34	532	12.3	2521	1	US-08-332-643-53	Sequence 53, Appl1
35	532	12.3	2521	1	US-08-332-638-59	Sequence 59, Appl1
36	509.5	11.8	2490	1	US-08-188-228-45	Sequence 45, Appl1
37	509.5	11.8	2490	1	US-08-332-628-45	Sequence 45, Appl1
38	481.5	11.1	3164	1	US-08-188-228-49	Sequence 49, Appl1
39	481.5	11.1	3164	1	US-08-332-638-43	Sequence 43, Appl1
40	481.5	11.1	3164	1	US-08-332-638-49	Sequence 49, Appl1
41	442.5	10.2	4104	1	US-07-998-003A-94	Sequence 94, Appl1
42	442.5	10.2	4104	1	US-08-453-274B-94	Sequence 94, Appl1
43	442.5	10.2	4104	1	US-08-453-695A-94	Sequence 94, Appl1
44	442.5	10.2	4104	1	US-08-268-161A-94	Sequence 94, Appl1
45	442.5	10.2	4104	2	US-08-453-702A-94	Sequence 94, Appl1

ALIGNMENTS

RESULT 1
US-08-431-560-2
Sequence 2, Application US/08431560
Patent No. 5620855
GENERAL INFORMATION:
APPLICANT: Anne H. Dantzig, et al.
TITLE OF INVENTION: Mammalian Influx Peptide
TITLE OF INVENTION: Transporter
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSER: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN
COUNTRY: U.S.A.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431,560
FILING DATE: 01-MAY-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/013,462
FILING DATE: 04-FEB-1993
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2499 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-431-560-2

Alignment Scores:

Pred. No.:	0	Length:	2499
Score:	4321.00	Matches:	832
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-10-025-380-1081 (1-832) x US-08-431-560-2 (1-2499)

```

Qy 1 MetIleuGlnAlaHisLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
Db 1 ATGATACTTCAGGCCCACTTCTTCACTCCCTGTGTCTTCTTATGCTTATTTGGCACTGGA 60
Qy 21 TyrGlyGlnGlyGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGlu 40
Db 61 TATGGCCAAAGGGGAAAGTTTAGTGGAACCCCTGAAACCATGACATTTTCTATTATGAA 120
Qy 41 GlyGlnGlyProSerGlnIleIlePheGlnPheLysAlaAsnProProAlaValThrPhe 60
Db 121 GGCCAGAACCCGAGTCAATATATTCAGTTTAAAGCCAACTCTCTGCTGTGACTTTT 180
Qy 61 GluLeuThrGlyGlyLysThrAspAsnIlePheValIleGluArgGlyGlyLeuLeuTyr 80
Db 181 GAACCTAAGTGGGAGACAGACACATATTTGTATAGAACGGGAGGACTTCTGTATTAC 240
Qy 81 AsnAlaGlyAlaLeuAspArgGlyLysThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
Db 241 AACGAGCCTTGAGCAGGAAACAAAGATCTACTCAATCTCCAGTTGACGCCCTGAC 300
Qy 101 AlaAsnGlyIleIleValGlnGlyProValProIleThrIleGluValLysAspIleAsn 120
Db 301 GCTATGGAATTAATAGAGAGGTCAGTCCCTATCACATAGAAAGAAAGACATCAAC 360
Qy 121 AspAsnArgProThrPheLeuGlnSerLysTyrGlyGlySerValArgGlnAsnSerArg 140
Db 361 GACATGACCCCAAGTCTTCTCAGTCAAGTACGAGGCTCAGTAAAGCAAACTCTGCC 420
Qy 141 ProGlyLysProPheLeuTyrValAlaAsnAlaThrAspLeuAspProAlaThrProAsn 160
Db 421 CCAGAAAGCCCTTCTGTATGTCAATGCCAGACCTGGATGATCCGGCCACTGCCAAT 480
Qy 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
Db 481 GGCCAGCTTATATCCAGATTTGTCATCCAGCTTCCATGATCAACATGTATGACTTT 540
Qy 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGlyGlySerGlnGlyLeuAsn 200
Db 541 CAGATCAACAAACAAAGGAGGACATCTCTTACCCGAGAGGATCTCAGAAATTGAT 600
Qy 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlyGlnSer 220
Db 601 CCTCTTAAGATCCCTTCTATATCTGTGTATCTCAGTGAAGGACATGGGAGGCGCAGAT 660
Qy 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValIleThrGlnAsnIleThrLys 240
Db 661 GAGATTCCTTCAGTATACCAATCTGTGATATCATAGTACAGAAATATTGGAAA 720
Qy 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
Db 721 GCACCAAAACCTGTGGAGATGTGTGAAAACTCAACTGATCTCAACCCCATCAAAATCACT 780
Qy 261 GlnValArgTyrPheAsnAspProGlyAlaGlnTyrSerLeuValAspLysGlyLysLeuPro 280
Db 781 CAGGTGGCGTGAATGATCCCGGTGCACAATATTCTTAACTTGAACAAAGAAACCTGCCA 840
Qy 281 ArgPheProPheSerIleAspGlnGlyLysAspIleTyrValIleThrGlnProLeuAspArg 300
Db 841 AGATTCCCATTTCAATGACCAAGAGAGATATTATCGTACCTCAGCCCTTGACCGA 900
Qy 301 GluGlyLysAspAlaTyrValPheTyrAlaValAlaLysAspGlyTyrGlyLysProLeu 320

```

```

Db 901 GAAGAAAAGAGTGCATATGTTTTTTATGCAAGTTGCAAGAGTACGAGAAACCACTT 960
Qy 321 SerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProThrCys 340
Db 961 TCATATCCGCTGGAATATTCGTAAAGATTAAGATTTATATATATATCCACTACATGTT 1020
Qy 341 ProSerProValThrValPheGlyValGlnGlyAsnGlyLysArgLeuGlyAsnSerIleGly 360
Db 1021 CCGTCAACAGTAAACCGTATTTAGAGTCCAGAGAAATGAACAGTGGGTAACTATCTCGG 1080
Qy 361 ThrLeuThrAlaHisAspArgAspGlyGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
Db 1081 ACCCTTACTCAGATGACAGGAGTGAAGAAATACCTGCCAAGCTTTCTTAACTACAGG 1140
Qy 381 IleValGlnGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
Db 1141 ATGTGGAGAACTCCCAAACTCCATGAGTGAAGTCTCTTCAATCAAACTTATGCT 1200
Qy 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyrAsnLeu 420
Db 1201 GGAATGTTACAGTTAGCTAAACAGTCTTGAGAAAGACAGATACCTCTCAGTCAACTTA 1260
Qy 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
Db 1261 ACGATAGAGGTGTCTGACAAAGATTTCAAGACCCCTTGTGTGCAAAATCAACGTTATT 1320
Qy 441 AspIleAsnAspGlnIleProIlePheGlyLysSerAspTyrGlyAsnLeuThrLeuAla 460
Db 1321 GATATCATATATAGATCCCATCTTTGAAAATTCAGATTATGAAACCTGACTCTTGCT 1380
Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
Db 1381 GAAGACAAACATTTGGGTCCACATCTTAAACATCCAGGCCACTGATGTGATGAGCCA 1440
Qy 481 PheThrGlySerSerLysIleLeuTyrHisIleIleGlyAspSerGlyArgLeu 500
Db 1441 TTTACTGGAGATTCTTAAATTTCTGTATCATATCATTAAGGAGACAGTGAAGCGCTG 1500
Qy 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
Db 1501 GGGGTGACACAAATCCCATACCAACCGGATATGTCTAATTAATAAGGCTCTTGAT 1560
Qy 521 PheGlnThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
Db 1561 TTTGAAACGACAGCTGTTTCCACATTTGTTCCAAAGAGAAATCTGAGCCTTACTG 1620
Qy 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValIleThrAspVal 560
Db 1621 TTTGTGTGAAGTACAAATGTCATGCTTTTGGCAAGTTCACGCTTATTTGTGACAGATGTG 1680
Qy 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
Db 1681 AATGAAGCACCTTAATTTTCCCAACAGTATTCAGAGCAAGAACTCAGTGAATGATGCT 1740
Qy 581 IleGlyThrLysValGlyAsnValIleThrAlaLysAspProGlyLysLeuAspIleSerTyr 600
Db 1741 ATAGGCACTTAATGTGGCAATGTGACTGCCAAGATCCAGAAAGTCTGGACATTAAGTAT 1800
Qy 601 SerLeuArgGlyAspThrArgGlyTyrPheLysIleAspHisValThrGlyLysIlePhe 620
Db 1801 TCACTGAGGGAGACCAAGAGGTTGGCTTAAATTAATGACACAGCTGCTGTGATGATCTT 1860
Qy 621 SerValAlaProLeuAspArgGlyAlaGlySerProTyrArgValGlnValAlaIleThr 640
Db 1861 AGTGTGCTTCATTTGACAGAGAGCGGAAAGTCCATATCGGTGATCAAGTGGGCCACA 1920
Qy 641 GluValGlyLysSerLeuSerSerValSerGlyPheHisIleLeuLeuMetAspVal 660
Db 1921 GAAGTAGGGGGTCTTCTTAACTCTGTCTCAGATTCACCTGATCTTATGATGATGTG 1980
Qy 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPheCysHisProLeu 680
Db 1981 AATGACAAACCTCCAGAGGTAGCCAAAGACTACACGGGCTTGTCTTCTGCAATCCCTTC 2040

```

QY 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspGlnHisLeuPheArgGly 700
 DB 2041 AGTCACCTGGAGAGCTCATTTTCGAGGCTACGATGATGATGACGACTTATTTTCGGGGT 2100
 QY 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnProGluValSerLys 720
 DB 2101 CCCCATTTTACATTTTCCCTCCGAGAGGAGGAGCTTACAAAACGACTGGAGAGTTTCCAAA 2160
 QY 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisPheGluGluValArgAlaTyr 740
 DB 2161 ATCAATGTAATCTATGCCCCGACTCTTACACAGGACACAGCTTTTGAGAGAGGCGCTAT 2220
 QY 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSerLeu 760
 DB 2221 GTGCTTGATCCGCATCAATGATGAGGGGTCGGCACCTTGAGAGGCACTTTTCTTTA 2280
 QY 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
 DB 2281 CCAATTACATTTCTGCACTGTGTGTGAGAGAGAGTTGTTCCGGCCAGAGGTCAACAGACT 2340
 QY 781 GluIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 DB 2341 GGGATACCACTGTGGGCAATGGAGGAGTGATGATGCTGACACCTTCTGTGTGATTGT 2400
 QY 801 IleIleLeuAlaValAlaPheIleArgIleLysLysAspLysGlyLysAspAsnValGlu 820
 DB 2401 ATATATTTTACAGTGTGTGTTTATCCGCAATTAAGAGAGATTAAGCAAAAGATATTTGAA 2460
 QY 821 SerAlaGlnAlaSerGluValLysProLeuArgSer 832
 DB 2461 AGTGCTCAAGCATCTGAAGTCAAACTCTGAGAAAGC 2496

RESULT 2

US-08-463-345-2
 ; Sequence 2, Application US/08463345
 ; Patent No. 5710018
 ; GENERAL INFORMATION:
 ; APPLICANT: Anne H. Pantzis, et al.
 ; TITLE OF INVENTION: Mammalian Influx Peptide
 ; TITLE OF INVENTION: Transporter
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Eli Lilly and Company
 ; STREET: Lilly Corporate Center
 ; CITY: Indianapolis
 ; STATE: IN
 ; COUNTRY: U.S.A.
 ; ZIP: 46285
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
 ; COMPUTER: Macintosh
 ; OPERATING SYSTEM: Macintosh
 ; SOFTWARE: Microsoft Word
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/463,345
 ; FILING DATE: 05-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/013,462
 ; FILING DATE: 04-FEB-1993
 ; INFORMATION FOR SEQ. ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2499 nucleotides
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA
 ; US-08-463-345-2

Alignment Scores: 0 Length: 2499
 Pred. No.: 4321.00 Matches: 832
 Score:

Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 1
 Gaps: 0

US-10-025-380-1081 (1-832) x US-08-463-345-2 (1-2499)

QY 1 MetIleLeuGlnAlaHisIleuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
 DB 1 ATATATCTTACAGGCCCATCTTCACTCCCTGTGCTCTTATGTTTATTTGGCACTGGA 60
 QY 21 TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGlu 40
 DB 61 TATGGCCAAGAGGGGAAAGTTAGTGACCCCTGAACCAATGACATTTCTATTATTAAGA 120
 QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheValAlaAsnProProAlaValIleThrPhe 60
 DB 121 GGCCAAAGAACCGAGTAAATTAATTAATTCAGTTTAAGGCCAATCTCTGTGTGACTTTT 180
 QY 61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluArgGlyLeuLeuTyrTyr 80
 DB 181 GAATTAATCTGGAGAGACAGACAAATATTTGTATGAACGGAGGACTTCTGTATTAAC 240
 QY 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 DB 241 AACAGAGCCTTGACAGAGGAAACAAAGATCTACTCAATCTCCAGGTGACGCTTGAC 300
 QY 101 AlaAsnGlyIleIleValAlaGlyLysProValProIleThrIleGluValLysAspIleAsn 120
 DB 301 GCTAATGGAATTAATTAATGAGGGGTCCAGTCCCTATCAATCAATGAATGAAGACATCAAC 360
 QY 121 AspAsnArgProThrPheLeuGlnSerLysTyrGluGlySerValArgGlnAsnSerArg 140
 DB 361 GACATATGACCCAGCTTTCTCCAGTCAAAAGTACGAAGCTCAGTAAGCAGAACTTCGC 420
 QY 141 ProGlyLysProPheLeuTyrValAlaAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
 DB 421 CCAAGAAAGCCCTTCTGTATGATGCATGACAGACCTGGATGATCCGGCACTCCCAAT 480
 QY 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
 DB 481 GGCCAGCTTATTAATCAAGATTCATCCAGCTTCCATGATCAATGATGATGATGATCTT 540
 QY 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlnSerGlnGluLeuAsn 200
 DB 541 CAGATCAACAAACCGGAGGCAATCTCTTACCCGAGGAGATCTCAGGAATGAAT 600
 QY 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnSer 220
 DB 601 CCGCTTAAGATCTCTCTATATATCGTGATCTCATGAAGACATGGAGGCCAGAGT 660
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValIleThrGluAsnIleTyrLys 240
 DB 661 GAGAAATTCCTTCAGTATACCAATCCATCTGGATATCATGTCACAGAAATATTTGGAAA 720
 QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
 DB 721 GACCAAAACCTGTGAGAGATGGTGAAGAACTCAACGATCTCACCCATCAAAATCACT 780
 QY 261 GlnValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGluLysLeuPro 280
 DB 781 CAGTGGGTGAGATGATCCCGGTGACAAATATCTTATGTTGACAAAGAGAGCTGCCA 840
 QY 281 ArgPheProPheSerIleAspGlnGluLysAspIleTyrValIleThrGlnProLeuAspArg 300
 DB 841 AGATTCCCATTTTCAATTAACCAAGAGAGATATTATGATCACTCAGCCCTTGACCCA 900
 QY 301 GluGluLysAspAlaIleTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu 320
 DB 901 GAGGAAAGAGATGATATGTTTATGCAAGTTGCAAGAGATGATGAGAAACCACTT 960
 QY 321 SerTyrProLeuGluIleHisValLysValLysAspIleAsnAspAsnProThrCys 340

Db 961 TCATATCCGCTGGAATTCATGTAAGATTAAGATTAATGATTAATCCACCTACTATGT 1020
 Qy 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 Db 1021 CCGCACACAGTAACCGTAATTTAGAGTCACAGAGATTAACAGACGAGGTAAACAGATCGGG 1080
 Qy 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
 Db 1081 ACCCTTACTGCACTGACAGGAGTAAAGAAATATCTCCAAACGTTTCTTAACTACAGG 1140
 Qy 381 IleValGlnGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 Db 1141 ATGTGTGAGCAAACTCCCAAACTTCCCATGGATGAGCTCTTCCAAATCCAAACCTTAATGT 1200
 Qy 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyrAsnLeu 420
 Db 1201 GGAATGTTACAGTTAGTAAACAGTCTCTGAAGAGAACGATCTCTCGATCACTTA 1260
 Qy 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
 Db 1261 ACGATAGAGGTGTGTGACAAAGATTTCAAGACCTTTGTTGCAAACTCAAGCTTAAT 1320
 Qy 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAla 460
 Db 1321 GATATCAATGATCATGATCCCATCTTGAATAATCAGATTATGGAACCTGACTTGTCT 1380
 Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db 1381 GAAGACCAACAACTGGGTCCACCATCTTAACTCAAGACGACCTGATGCGATGAACCA 1440
 Qy 481 PheThrGlySerSerLysIleLeuTyrHisIleIleLysGlyAspSerGluGlyArgLeu 500
 Db 1441 TTTTACTGGAGCTTAAATTTCTTATCATATCAATTAAGAGACAGATGAGGAGCGCTG 1500
 Qy 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
 Db 1501 GGGGTGACACAGATCCCATACCAACACCGGATATGTCAATTAATAAAGCCCTTGAT 1560
 Qy 521 PheGluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 Db 1561 TTTGAAACAGACGCTGTTTCCAACTGTGTTCAAGACGAAATCCCTGAGCCCTGATGTG 1620
 Qy 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 Db 1621 TTTGGTGTGAGTCAATGCAATGCAATGCTTTTCCCAAGTTCAAGCTTATTTGACAGATGTG 1680
 Qy 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 Db 1681 AATGACCACTCAATTTTCCCAACAGTATTTCCAAAGCAAGTCAAGTGAAGATGTAGCT 1740
 Qy 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTyr 600
 Db 1741 ATAGGCACTAAAGTGGGCAATGTATCTCCAAAGATCCAGAAAGTCTGGAATATAGCTAT 1800
 Qy 601 SerLeuArgGlyAspThrArgGlyTyrPheLysIleAspHisValThrGlyGluIlePhe 620
 Db 1801 TCACCTAGGGAGACACAGAGGTTGGCTTAAATTTGACCAACGTCGTCGAGATCTTT 1860
 Qy 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaAlaThr 640
 Db 1861 AGTGTGCTCATTTGACAGAGAGAGCCGGAAGTCCATTCGGGTATCAAGTGTGGCCACA 1920
 Qy 641 GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 Db 1921 GAAGTAAAGGGGGGTCTTCTTAAAGCTGTGTGTCAAGTTCACCTGATCTTATGAAATGTG 1980
 Qy 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 1981 AATGACCACTCCCAAGCTTACCAAGACATCAACAGGCTGTTCTTGTGCAATCCCTTC 2040
 Qy 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheArgGly 700
 Db 2041 AGTCACTGGAAGTCTATTTTGAAGGCTACTGATGATGATCAAGCACTTATTTTGGGGGT 2100

Qy 701 ProIlePheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerLys 720
 Db 2101 CCCCATTTTACATTTTCCCTCGGCATGAGCTTACAAAGACATGGGAAAGTTTCCAA 2160
 Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaTyr 740
 Db 2161 ATCAATAGTACTCATCCCGACATGCTGTACAGGACACAGACTTGTAGAGAGGGGGTAT 2220
 Qy 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGlnGluIleValSerLeu 760
 Db 2221 GTCTGTTATCCGATCAATGATGAGGGGTCGGCCACCTTGGAGAGCATTTGTTCTTAA 2280
 Qy 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
 Db 2281 CCAGTACATCTTCGACGTTGTGTGAGAGAAAGTGTTCGGGCACAGCTGACCAAGCT 2340
 Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2341 GGGATACCACTGTGGGACATGACATGGTATTAATCTGACCAACCTTCTGTGATTTGT 2400
 Qy 801 IleIleLeuAlaValAlaPheIleArgIleLysLysAspLysGlyLysAspAsnValGlu 820
 Db 2401 ATATATTTAGCACTTGTGTATCCGATTAAGAGATTAAGGCAAGATTAATGTTGAA 2460
 Qy 821 SerAlaGlnAlaSerGluValLysProLeuArgSer 832
 Db 2461 AGTCTCAAGCATCTGAAGTCAAACTCTGAGAGC 2496

RESULT 3
 US-08-188-228-47
 Sequence 47, Application US/08188228
 Patent No. 5597725
 GENERAL INFORMATION:
 APPLICANT: Suzuki, Shintaro
 TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
 NUMBER OF SEQUENCES: 62
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ADDRESSER: Borun
 STREET: 6300 Sears Tower, 233 S. Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/188,228
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/049,460
 FILING DATE: 19 APR 1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/872,643
 FILING DATE: 17 APR 1992
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5597725and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 31340
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELEX: 25-3856
 INFORMATION FOR SEO ID NO: 47:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3048 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-188-228-47

Alignment Scores:

Pred. No.:	4,4e-65	Length:	3048
Score:	684.50	Matches:	176
Percent Similarity:	46.92%	Conservative:	106
Best Local Similarity:	29.28%	Mismatches:	288
Query Match:	15.84%	Indels:	31
DB:	1	Gaps:	12

US-10-025-380-1081 (1-832) x US-08-188-228-47 (1-3048)

```

QY      239  TrrpysalAprolysProvalGluMetValGluasnSerThrAspProHisProIleLys 258
DB      525  TGGGTATCCCA---CCCATCAAGCTGCCGAGAACTCGCGGGGCTTCCCGGCACAG 581
QY      259  IleThrGlnValAArgTrrp-----AsnAspProGlnAlaGlnTyrSerLeu----- 273
DB      582  CTCGTGAGATCCCGTCCGACAAAGACATGACATCCCATCCGCTACAGATCAGCGGA 641
QY      274  ValAspLysGluLysLeuProAlaArgPheProPheSerIleAsp---GlnGluGlyAspIle 292
DB      642  GTGGGTCCGACACGCCCCCATGAGAGTCTTCAGCATTAATCTCATGTCGCGCGGATG 701
QY      293  TyrValThrGlnProLeuAspArgGluGlnLysAspAlaTyrValPheTyrAlaValAla 312
DB      702  TAGGTCAAGAGGCCCATGAGACCGGAGAGACCCCTTACCATCCATCCGACGCCACGCT 761
QY      313  LysAspGluTyrGlyLysProLeuSerTyrProLeuGlnIleHisValLysValLysAsp 332
DB      762  GTGACATGATGATGCACCAAGGTGAGAACCCCTACCATCCGCTACATCTACATCATGAC 821
QY      333  IleAsnAspAsnProProThrCysProSerProValThrValPheGluValGlnGluAsn 352
DB      822  ATGATATACAAACCACTGAGATTATCAACAGGTGTTCACAACTGCTCCGTGACGAGGCG 881
QY      353  GluArgLeuGlnLysAsnSerIleGlyThrLeuThrAlaHisAspArgAspGlnGluAsnThr 372
DB      882  TTCAGAGCAGGACCACTACGTGATGATGATGATGATGATGATGATGATGATGATGATG 941
QY      373  AlaAsnSerPheLeuAsnTyrArgIleValGlnGlnThrProLysLeuProMetAspGly 392
DB      942  GCCAAGCGGATGCGGTACCGGATCGGTACCCAGACCCACAGACCCGCTGCCAGAAAT 1001
QY      393  LeuPheLeuIleGlnThrTyrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLys 412
DB      1002  ATGTTTACCATCAACAGCAGAGATGAGATATGCTACAGTGGCGGCTGGCTGGAGCGA 1061
QY      413  GlnAspThrProGlnTyrAsnLeuThrIleGluValSerAspLysAsp----- 428
DB      1062  GAGAAATTCAGAGTACAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1121
QY      429  -----PheLysThrLeuCysPheValGlnIleAsnValIleAspIleAsnAspGlnIle 446
DB      1122  TATGGCTCTCAAAACACAGCCACAGCATCATCAGGTGACAGATGATGATGATGATGATG 1181
QY      447  ProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAlaGluAspThrAsnIleGly 466
DB      1182  TCAGAAATTTACCGCCAGCAGCTTTCAGAGG-----GAGGTCCCGCAAAACAGGTGAG 1235
QY      467  SerThrIleLeuThrIleGlnAlaThrAspAlaAspGluProPheThrGlySerSerLys 486
DB      1236  ACCGTGTCCGAAACCTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATG 1295
QY      487  IleLeuTyrHisIleIleLysGlyAspSerGlnGlyArgLeuGlyLysAspThrAspPro 506
DB      1296  GCCGTTTACCGCATCATCAGTGGGATCCATCCGCGGACATTCAGCGCCGACAGACCC 1355
QY      507  HisThrAspThrGlyTyrValIleIleLysLysProLeuAspPheGlnThrAlaAlaVal 526
DB      1356  GTAAACCAACGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1415

```

```

QY      527  SerAsnIleValPheLysAlaGluAsnProGluProLeuValPheGlyValLysTyrAsn 546
DB      1416  TTCATGCTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1475
QY      547  AlaSerSerPheAlaLysPheThrLeuIleValThrAspValAsnGluAlaProGlnPhe 566
DB      1476  TTCAGTCCAGCGCAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1535
QY      567  SerGlnHisValPheGlnAlaLysValSerGlnAspValAlaIleGlyThrLysValGly 586
DB      1536  CCTCAAAACCAAGCTGATCCGCTGAGAGAGGCGGTGCCCCCGGACCGTGTAC 1595
QY      587  AsnValThrAlaLysAspProGlnGlyLeu-----AspIleSerTyrSerLeuArg 603
DB      1596  ACCTTTTACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1655
QY      604  GlyAspThrArgGlyTyrLeuLysIleAspHisValThrGlyLysIlePheSerValAla 623
DB      1656  TCAGAACCCAGGAGCTGCTGACATCAATGACCAACAGCGGACATCACACGAGTGGA 1715
QY      624  ProLeuAspArgGlnAla-----GlySerProTyrArgValGlnValAla 639
DB      1716  GTCTGACCGGTAGTCCCTTACACCAAAACAGCTTACAGAGCCACTTCTGGA 1775
QY      640  ThrGluValGlyLysSerSerSerSerSerValSerGluPheHisLeuIleLeuMetAsp 659
DB      1776  GCTGACAAATGGATATACCCCGGACGCGGACCGGACCGGACCGGACCGGACCGGACCG 1835
QY      660  ValAsnAspAsnProProAlaGluAlaLysAspTyrThrGlyLeuPhePheCysHisPro 679
DB      1836  ATCAACGACAAACCCCTGAGTGGCCCAAGAGGCGGACGATGCTGCGAGAGGCCAAC 1895
QY      680  LeuSerAlaProGlySerLeuIlePheGlnAlaThrAspAspArgGlnHisLeuPheArg 699
DB      1896  CTGAGAGCC-----ATCAACATCAGCGCGCGGACGCTGACGTGACGCCAACATC 1946
QY      700  GlyProHis---PheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTrrpGluVal 718
DB      1947  GGCCTTACGCTTTCAGAGCTGCCCTTTCGCCGCGGCGCGGCGGCGGAGAACTGACATC 2006
QY      719  SerLysIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGlnGluArg 728
DB      2007  ACCGCTGAAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2066
QY      739  AlaTyrValValLeuIleArgIleAsnAspGlyLysArgProProLeuGlnGlyIleVal 758
DB      2067  ATGATGACGCTCCCATCATCTGACAGACTTGGAAACCTTCCCTGTCMAACAGCTCC 2126
QY      759  SerLeuProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHis 778
DB      2127  ATATCAAGTCAAGGTGTGCCATGTATGATGATGATGATGATGATGATGATGATGATGATG 2183
QY      779  GlnThrGlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuVal 798
DB      2184  -----GCAGTGACAGCGGTGCTGGGACCGGATGATGATGATGATGATGATGATGATG 2237
QY      799  IleGlyIleIleLeuAlaValAla-----PheIleArgIleLysLysAspLysGlyLys 816
DB      2238  ATCTCATCTGCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2297
QY      817  Asp 817
DB      2298  GAG 2300

```

RESULT 4

US-08-332-643-41
Sequence 41, Application us/08332643
Patent No. 5639634
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
NUMBER OF SEQUENCES: 56

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ADDRESSEE: Bicknell
 STREET: Two First National Plaza, 20 South Clark
 STREET: Street
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60603
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/332,643
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/872,643
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5639634and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 27866/30795
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 346-5750
 TELEFAX: (312) 984-9740
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 41:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3048 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-332-643-41
 Alignment Scores:
 Pred. No.: 4.4e-65 Length: 3048
 Score: 684.50 Matches: 176
 Percent Similarity: 46.92% Conservative: 106
 Best Local Similarity: 29.28% Mismatches: 288
 Query Match: 15.84% Indels: 31
 Gaps: 12
 US-10-025-380-1081 (1-832) x US-08-332-643-41 (1-3048)
 QY 239 TrpysalAProlysaProValGluMetValGluAsnSerThrAspProHisProIleVal 258
 Db 525 TGGGTGATCCCA---CCCATCAAGTCGCCCGAAGAACTGGCGGGCCCTTCCCGGAGAG 581
 QY 259 IleThrGlnValArgTyr-----AsnAspProGlnAlaGlnTyrSerLeu----- 273
 Db 582 CTCGTGAGGATCCGTCGACCAAGACATGATCAATCCCATCCGATACAGATCAAGGGA 641
 QY 274 ValAspIlysgIuIy/sleuProArgPheProPheSerIleAsp---GlnGluGlyAspIle 292
 Db 642 GTGGGTGCCGACAGCCGCCCATGAGAGTCTTCAAGATTAACTCCATGTCGGCGGAGTG 701
 QY 293 TyrValThrGlnProLeuAspArgGluGluIuIyAspAlaTyrValPheTyrAlaValAla 312
 Db 702 TACGTCAAGAGCCCATGATGACCGGAGAGACAGCGCTTTACCACTCCGAGCCCAAGCT 761
 QY 313 LysAspGluTyrGlyIy/sProLeuSerTyrProLeuGluIleHisValIy/sValIyAsp 332
 Db 762 GTGACATGATGAGCAACAGATGAGAGAACCCATCGACATGATCATGATCATGATGATG 821
 QY 333 IleAsnAspAsnProProThrCysProSerProValThrValPheGluValGlnGluAsn 352
 Db 822 ATGATATGACCAACCACTCGAGTTTCATCAACAGGCTTCAACATGCTCGTGAAGAGGC 881
 QY 353 GluArgLeuGlyAsnSerIleGlyThrLeuThrAlaHisAspArgAspGluGluAsnThr 372

Db 882 TCACAGCAGCAGCACTACGATGATGATGATGATGATGATGATGATGATGATGATGATG 941
 QY 373 AlaAsnSerPheLeuAsnTyrArgIleValGluGlnThrProIy/sLeuProMetAspGly 392
 Db 942 GCCACGCGAGTGGTGGTATCGGATCGTGAACCCACCAAGCCCGCTCCAGAAAT 1001
 QY 393 LeuPheLeuIleGlnThrTyrAlaGlyMetLeuGlnLeuAlaIy/sGlnSerLeuIy/s 412
 Db 1002 ATGTTCAACCTCAACAGCAGATGATGATGATGATGATGATGATGATGATGATGATG 1061
 QY 413 GlnAspThrProGlnTyrAsnLeuThrIleGluValSerAspIy/sAsp----- 428
 Db 1062 GAGAAAGTTACAGATACACATGATGATGATGATGATGATGATGATGATGATGATGATG 1121
 QY 429 -----PheIy/sThrLeuIy/sPheValGlnIleAsnValIleAspIleAsnAspGlnIle 446
 Db 1122 TATGGCTCTCAACACAGCCACAGCCATCATCAGGTGACAGATGATGATGATGATGATG 1181
 QY 447 ProIlePheGluIy/sSerAspTyrGlyAsnLeuThrLeuAlaGluAspThrAsnIleGly 466
 Db 1182 TCAGATTTTACCGCCAGCAGCTTTCAGAGG-----GAGTCCCGCAAAACAGCTGGAG 1235
 QY 467 SerThrIleLeuThrIleGlnAlaThrAspAlaAspGluProPheThrGlySerSerIy/s 486
 Db 1236 ACCGTGTCGCAAACTCAACGATGATGATGATGATGATGATGATGATGATGATGATGATG 1295
 QY 487 IleLeuTyrHisIleIleIy/sGlyAspSerGluGlyArgLeuIy/sValAspThrAspPro 506
 Db 1296 GCCGTTTACCGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1355
 QY 507 HisThrAsnThrGlyTyrValIleIleIy/sIy/sPheProLeuAspPheGluThrAlaAlaVal 526
 Db 1356 GTTACACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1415
 QY 527 SerAsnIleValPheIy/sAlaGluAsnProGluProIleValPheGlyValIy/sTyrAsn 546
 Db 1416 TTCATGCTGACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1475
 QY 547 AlaSerSerPheAlaIy/sPheThrIleIleValIleThrAspAlaGlnGluIleAspGlnPhe 566
 Db 1476 TTCATGCTGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1535
 QY 567 SerGlnHisValPheGlnAlaIy/sValSerGluAspValAlaIleGlyThrIy/sValGly 586
 Db 1536 CCTCAAAACCAAGCTGATCCGCTGAGAGAGGCGTCCCGCCCGCAGCTGCTGATGATG 1595
 QY 587 AsnValThrAlaIy/sAspProGluIy/sLeu-----AspIleSerTyrSerLeuArg 603
 Db 1596 ACCTTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1655
 QY 604 GlyAspThrArgIy/sIy/sIleAspHisValThrIy/sGluIlePheSerValAla 623
 Db 1656 TCAGACCCAGCAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1715
 QY 624 ProLeuAspArgGluAla-----GlySerProTyrArgValGlnValAla 639
 Db 1716 GTGTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1775
 QY 640 ThrGluValGlyIy/sSerSerLeuSerSerValSerGluPheHisIleuIleMetAsp 659
 Db 1776 GTGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1835
 QY 660 ValAsnAspAsnProProIleGluAlaIy/sAspTyrThrIy/sLeuPhePheCysHisPro 679
 Db 1836 ATCAACGACAAACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1895
 QY 680 LeuSerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisIleuPheArg 699
 Db 1896 CTGAAGCC-----ATCAACATTCAGGCGGCGGCTGATGATGATGATGATGATGATG 1946
 QY 700 GlyProHis---PheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluVal 718

Db 1947 GGCCCTACGCTTCGAGCTGCCCTTTGTCGGGCGCGCCGGAAGAACTGACCATC 2006
 QY 719 SerlyslleanglyThrhlealargleuserThrhghisthrasphlegluarg 738
 Db 2007 ACCGCTGAACGGTGAAGTATGCGCACTGAGCTTGGGCACTGCTGACGAGCGCGG 2066
 QY 739 AlATyValValleulleargyleasnaspglyglargpropleuglylval 758
 Db 2067 ATGATGACGCTCCCATCATGCTGACAGACTCTGGAACCTCCCTGTCACACGCTCC 2126
 QY 759 SerleueprovalThrhiecyserCyserValgluglyserCyshreagprolaaglyhis 778
 Db 2127 ATCTCAAGTCAAGTGTGCGCATGTGATGACAAACGGGACCTGACCATTCCTATGTC 2183
 QY 779 GlnthrylleproThrhValglYmetAlaValglYlleleuleuthrhThrhleuVal 798
 Db 2184 -----GCAGTGGAGCGGCGCTGCTGCGGACCGGTGCGCATCGGCCATCTCATGTC 2237
 QY 799 lileglylleleuleuAlaVal-----Pheileargyleuslyaspglylys 816
 Db 2238 ATCTCATCTGCTGACCATGCTCTGCTGTTTGTGATGATGAGCGGCGAGAGAG 2297
 QY 817 Asp 817
 Db 2298 GAG 2300

RESULT 5

US-08-332-638-47
 Sequence 47 Application US/08332638
 Patent No. 5646250
 GENERAL INFORMATION:
 APPLICANT: Suzuki, Shintaro
 TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
 NUMBER OF SEQUENCES: 62
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 STREET: 6300 Sears Tower, 233 S. Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/332,638
 FILING DATE: 01-NOV-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/872,643
 FILING DATE: 17 APR 1992
 APPLICATION NUMBER: US/08/049,460
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5646250and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 31340
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 47:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3048 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-332-638-47

Alignment Scores:

Pred. No.: 4,4e-65 Length: 3048
 Score: 684.50 Matches: 176
 Percent Similarity: 46.92% Conservative: 106
 Best Local Similarity: 29.28% Mismatches: 288
 Query Match: 15.84% Indels: 31
 DB: 1 Gaps: 12

US-10-025-380-1081 (1-832) x US-08-332-638-47 (1-3048)

QY 239 TTPlyAlaProlyspProValglumetValgluasnerThrhspProhispProlely 258
 Db 525 TGGGTCAATCCA---CCCATCAAGTGCCTCGAAGAACTCGGCGCGCTTCGCGACAG 581
 QY 259 lileThrhValArgTTP-----AsnaspglylValgluThrhserleu----- 273
 Db 582 CTGTGAGATCCGCTGTCGCAAGAACATGACATCCCATCCGATACAGATCAGCGGA 641
 QY 274 ValAspglyglulysleupProArgPheProPheSerlleasp--Gingluglyaspile 292
 Db 642 GTGGGTGCGCAGCAGCGCCCATGAGAGGTCTTCAAGATTATCATGTCGCGCGGATG 701
 QY 293 TyrValThrhGlnProleuaspaTggluglulysaspAlaTyrValPheTyrAlaValAla 312
 Db 702 TACGTCAAGAGCGCCATGAGCCGAGAGACAGCCCTTATCAACCTCCAGGCCACGCT 761
 QY 313 LyAspglylurTgyllysProleuserTyrProleugluThrhVallyAsp 332
 Db 762 GTGACATGATATGCAACAAAGTGAAGAACCCATGACCTGTACATCTACATGATGAC 821
 QY 333 lileAsnaspaPProProThrhCyserProserProValThrhValPhegluValgluasn 352
 Db 822 ATGAATGACAAACACACCTGATTCATCAACCAAGTGTCAACAGTCTCCGTGAGAGAGGC 881
 QY 353 GluArgleuglyasnerlleglyThrhleuThrhAlaHisAspArgspgluasThr 372
 Db 882 TCCAAAGCCAGCACTTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 941
 QY 373 AlaAspserPheleuasnTyrArglleValglugluThrhProlyleupPromeleAspgly 392
 Db 942 GCCAAGCGGATGTGCGGTACCGGATCGTACCCAGACCCCAAGACCCGCTCCAGAAAT 1001
 QY 393 leupheleulleGlnThrhTyrAlaGlymetleugluThrhleuAlaYsglnserleuLys 412
 Db 1002 ATGTTACATCAACAGCGAGCATGAGATATGTCACAGTGGCGGTGCTGAGACCA 1061
 QY 413 GluAspThrhProGlnTyrAsnleuthrhleGluValSerAspLysasp----- 428
 Db 1062 GAGAAAGTTCAAGCATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1121
 QY 429 -----PheIysThrhleuCyshreValGlnlleAsnVallleAspilleAsnAspGlnlle 446
 Db 1122 TATGCGCTCTCAAAACAGCCACAGCCATCATGATGATGATGATGATGATGATGATGAT 1181
 QY 447 ProillePheglulysSerAspTyrGlyAsnleuthrhleuAlaGluAspThrhAsnillegly 466
 Db 1182 TCAGAAATTTACCGCAGCAAGCTTGAGG-----GAGGTCCCGAAGAACAGCGTGAAG 1235
 QY 467 SerThrhilleuThrhilleGlnAlaThrhAspAlaAspGluProPheThrhGlySerLys 486
 Db 1236 ACCGTGATCGCAAACTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1295
 QY 487 lileleuThrhillellelysglyAspsegluglyArgleuglyValAspTyrAspPro 506
 Db 1296 GCCGTTTACGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1355
 QY 507 HisThrhAsnThrhTyrValillellelyslYsPProleuasPhegluThrhAlaVal 526
 Db 1356 GTAAACCAAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1415
 QY 527 SerAsnilleValPheylsAlaGluAsnProGluProleuValPhegluVallyslYsP 546
 Db 1416 TTCATGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1475

```

Qy 547 AlasSerPheAlaValPheThrLeuValThrAspValaGluAlaProGlnPhe 566
Db 1476 TTCAGTTCACGGAGGGGTGACATCTCCATGACATGACAGAGGCTCTACTTC 1535
Qy 567 SerGlnHisValPheGlnAlaValSerGlnAspValAlaIleGlyThrValGly 586
Db 1536 CCCCAACACCAACAGCTGATCCGCTGGAGAGGGCGGCCCGGACCGGTCTAC 1595
Qy 587 AsnValThrAlaValAspProGlnGlyLeu-----AspIleSerTyrSerLeuArg 603
Db 1596 ACCTTTTCAGCTGGAGACCCGAGCCGTTTCATGACAGAGGCTGGAGATCTCAAACTG 1655
Qy 604 GlyAspThrArgGlyTyrPheLeuValIleAspHisValThrGlnIlePheSerValAla 623
Db 1656 TCACAGCCCGAGAGCTGGCTGACATGACATGACCAAGAGAGATCCAGCTGGCA 1715
Qy 624 ProLeuAspArgGluAla-----GlySerProTyrArgValGlnValAla 639
Db 1716 GTGCTGGACCGTGAAGTCCCTTACACCAAAACAGCTACGAGGACACTTCTGCA 1775
Qy 640 ThrGluValGlyGlySerSerLeuSerValSerGlnPheHisLeuIleMetAsp 659
Db 1776 GCTGACATGGAGATACCCCGGACCGGACCGGACCCCTCAAGATCTATCTCATTTGAC 1835
Qy 660 ValAsnAspAsnProProArgLeuAlaValAspTyrThrGlyLeuPheCysHisPro 679
Db 1836 ATCAACAGCAACGCCCTGAGCTGCTGCCAAGAGGCGCAGATCTGGACAGGCCCAAC 1895
Qy 680 LeuSerAlaProGlySerLeuIlePheGlnAlaThrAspAspArgGlnHisLeuPheArg 699
Db 1896 CTGAACGCC-----ATCAACATCACGGCGCGCCAGCGCTGACGATGACCCCAACATC 1946
Qy 700 GlyProHis---PheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluVal 718
Db 1947 GGCCCTCAACGCTTCGAGCTGCTCCCTTGTCCCGCGCGCGGAGAAAGACTGACCATC 2006
Qy 719 SerValIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluValArg 738
Db 2007 ACCCGCTGACAGCGTACTATGCCCCAACTCACTGCGCATCTGATCTGAGAGCCGGG 2066
Qy 739 AlaTyrValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGlnGlyIleVal 758
Db 2067 ATGATGACGTCCTCCATCATGTCACAGACTCTGGAACCTCCCTGTTCCACACGCTCC 2126
Qy 759 SerLeuProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHis 778
Db 2127 ATCTCAAAAGTCAGGTGTGCCCATGTGATGACAAACGGGAGCTGACACCATTTGGC-- 2183
Qy 779 GlnThrGlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuVal 798
Db 2184 -----GAGTGGACAGCGGCTGGTGTGGGACCGGTGCCATGCGCATCTCATCTGCC 2237
Qy 799 IleGlyIleIleLeuAlaValIle-----PheIleArgIleIleValAspGlyGly 816
Db 2238 ATCTCATCTCCGTCGACCATGCTCTGTTTTCATGAGTGAAGAGCGGAGAGAG 2297
Qy 817 Asp 817
Db 2298 GAG 2300

```

RESULT 6
US-09-643-597-126
Sequence 126, Application US/09643597
Patent No. 6426072
GENERAL INFORMATION:

APPLICANT: Wang, Tongtong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hoeken, Nancy
APPLICANT: Fanger, Gary R.
APPLICANT: Li, Samuel X.

```

APPLICANT: Wang, Atjun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.455C11
CURRENT APPLICATION NUMBER: US/09/643,597
NUMBER OF SEQ ID NOS: 369
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 126
LENGTH: 3552
TYPE: DNA
ORGANISM: Homo sapien
US-09-643-597-126

```

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
1	550	63	48.27%	15.54%	3552	187	106	263	51	22

US-10-025-380-1081 (1-832) x US-09-643-597-126 (1-3552)

```

Qy 239 TyrPheAlaProIleAspProValGluMetValGluAsnSerThrAspProHisProIleVal 258
Db 476 TGG---GCACCTTTCCTTCTGCTATGACAGAAATTCCTTGGCCCTTCCATTGTTT 532
Qy 259 IleThrGlnValArgTyrPheAspProGlyAlaGlnTyrSerLeu----- 273
Db 533 CTTCACACAGTTGAA---TCTGATGACAGCAGACAACTATATCTCTTACTGATTAAGT 539
Qy 274 -----ValAspGlyGlyLeuProArgPheProPheSerIleAspGlnGly 290
Db 590 GAGCTGAGACTGATTAAGAACTTAATTTGTTTATATGAAGAGAC--ACTGGA 646
Qy 291 AspIleTyrValThrGlnProLeuAspArgGluValAspAlaTyr-----ValPhe 308
Db 647 AATCTATTGACCTCGCGCTGTGATCTGAAGATATGATTTTGTGATTGATTGCT 706
Qy 309 TyrAlaValAlaValAspGluTyrGlyLeuProLeuSerTyrProLeuGlnIleHisVal 328
Db 707 TATGCTCACTCACTGATGATATTCAGCAGATCTG-----CCCTCCCATCACTCCAT 760
Qy 329 LysValIleAspIleAsnAspAsnProProThrCysProSerProValThrValPheGln 348
Db 761 AGGTGAGAGATGAATAATGACACACACCTGTTTTCACAGAACAAATTTTAATTTGAA 820
Qy 349 ValGlnGluAsnGluArgLeuGlyAsnSerIleGlyThrLeuThrAlaHisAspArgAsp 368
Db 821 GTTTTGGAAAGTGTGAGACTGGTACTACAGTGGGGGGGTTTGTGCCACAGCAGAT 880
Qy 369 GlnGluAsnThrAlaAsnSerPheLeuAsnTyrArgIleValGlnGlnThrProIleVal 388
Db 881 GAACCGACACATGATGATCGCCCTGAATATACGATTTTTCAGCAGACACCAAGTCA 940
Qy 389 ProMetAspGlyLeuPheLeuIleGlnThrTyrAlaGlyMetLeuGlnLeuAlaValGln 408
Db 941 CCT-----GGGCTCTTTTGTGATCCAGACAGGCGGTATACACACAGTCTGAT 994
Qy 409 SerLeuValSerGlnAspThrProGlnTyrAsnLeuThrIleGluValSerAspValAsp 428
Db 995 TATTTGACAGAGAGGTGTGAGCAAGTACTTATATGAAGTACAAAGCATGAGAT 1054
Qy 429 PheLys-----ThrLeuCysPheValGlnIleAsnValIleAsp 441
Db 1055 GGCAGGTTTGTGATGATGACATCACTTGTATC-----ATAACGTTAACAGAT 1108
Qy 442 IleAsnAspGlnIleProIlePheGlnIleValSerAspTyrGlyValAsnLeuThrLeuAlaGln 461

```

```

Db      1109 TCAATATATATGACCCACCTTTCAGACAAAATGCTTATGAA-----GCATTGTGATAG 1162
QY      462 AspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluProPhe 481
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1163 GAAATATGATTCATTCATGGAATATCTTACGAAATACCTTACGAAATGAAAGATTTTAATTAAAC 1222
QY      482 ThrGlySerSerIleLeuThrIleIleLeuGlyAspSerGluGlyArgLeuGly 501
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1223 ACTGCCAATGAGAGTCAATTTTACATTTTAAAGGAAATGAAATGACATTTCAAA 1282
QY      502 ValAspThrAspProHisThrAsnThrGlyTyrValIleIleLeuGlySerProLeuAspPhe 521
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1283 ATCGACACAGCAAAAGAACTAAAGAAAGTCTTCTGTTGTAAGCCACTGAATAT 1342
QY      522 GluThrAlaAlaValSerAsnIleValPheGlyAlaGluAsnProGluProLeuValPhe 541
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1343 GAAGAAACCGTCAGTGAAGAACCTGGAATTCGAAATGCAATGAAGCCACTTGTCTAGA 1402
QY      542 GlyVal--LysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1403 GATATTCACAGAGTACAGCCTTGACAGAGCCTTGTTACAGTTCATGTAGAGGATCTG 1462
QY      561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1463 GATGAGGAGCCTGATGACACTCTGACAGCCCAATATGTGCGGATTTAAAGAACTTGA 1522
QY      581 IleGlyThrIleValGlyAsnValThrAlaLysAspProGlu-----GlyLeuAsp 597
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1523 GTGGGGTCAAAGATCAACGGCTATAGGCATATATACCCCAAAATAGAAATGGCAATGCT 1582
QY      598 IleSerTyrSerLeuArgGlyAspThrArgGlyTyrLeuLysIleAspHisValThrGly 617
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1583 TTAAAGTACAAAATAATGATGATCTTAAGGTTGATCACCATTTGATGAATTTACGGG 1642
QY      618 GluIlePheSerValAlaProLeuAspArgGluAlaGlySerPro-----Tyr 633
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1643 TCAATCTAACTTCCAAATCTCGATGAGGAGGTTGAACTCCCAAAAATGAGCTGTAT 1702
QY      634 ArgValGlnValAlaThrGluValGlySerSerLeuSerSerValSerGluPhe 653
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1703 AATATTACATCTCTCGGCAATAGACAAAGATGATGATGATCTGGAACACTTGTCTGTG 1762
QY      654 HisLeuIleLeuMetAspValAsnAspAsnProProArgLeuAlaLysAspTyrThrGly 673
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1763 AAC-----ATTGAAGATGTAATGATATCCACAAATACTTCAAGATAT----- 1810
QY      674 LeuPhePheCysHisProLeuSerAlaProGlySerIleLeuGluAlaThrAspAsp 693
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1811 GTAGTCATTTGCAAAACCAAAATGGGGTATACCGACATTTTA-----GCTGTTGATCT 1864
QY      694 AspGlnHisLeuPheArgGlyProHisPheThrPheSerIleGlySerGlySer----- 711
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1865 GATGAACCTGTCATGAGAGCTCA--TTTATTTTCAGTTTGCCCAATCTTCCCGAA 1921
QY      712 LeuGlnAsnAspTyrGluValSerIleLysGlyThrHisAlaArgLeuSer---Thr 730
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1922 ATCAGTAGACTGTGAGAGCTCACCAAGTTAATGATACAGCTGCCGCTTTCATATCAG 1981
QY      731 ArgHisThrAspPheGluGluArgAlaTyrValIleLeuIleArgIleAsnAspGlyGly 750
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1982 AAAAATGCTGATTTCAAGAA-----TATACCATTCCTATTCTGTAAGAAAGAC----- 2029
QY      751 ArgProProLeuGluGlyIleValSerLeuProValThrPheCysSerCysValGluGly 770
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2030 AGGGCCGGCCAGCTGCAACAAATTAATTAAGATTAATCTGTGTGATGATCTCATCA 2089
QY      771 SerCysPheArgProAlaGlyHisGlnThrGlyIleProThrValGlyMetAlaVal--- 789
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2090 ACTCAGTGTGTGAGATTCAGAGAGTACAGAGTATACTTGGAAAAATGGGCAATCTT 2149
QY      790 GlyIleLeuLeuThrThrLeuLeuValIleGlyIleIleLeuAlaValPheIleArg 809
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2150 GCAATATTACTGGGTATAGCACTGCTCTTTCTGATATGCTAACTTAGATATGAGATT 2209

```

```

QY      810 IleLysLysAspLysGlyLys 816
Db      2210 TTTGTGCACTTAAGGAAA 2230

RESULT 7
US-08-188-228-61
; Sequence 61, Application US/08188228
; Patent No. 5597725
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,228
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/049,460
; FILING DATE: 19 Apr 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/872,643
; FILING DATE: 17 Apr 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5597725and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31340
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2690 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-188-228-61

Alignment Scores:
Pred. No.: 1,59e-62 Length: 2690
Score: 660.50 Matches: 205
Percent Similarity: 42.31% Conservative: 103
Best Local Similarity: 28.16% Mismatches: 288
Query Match: 15.29% Indels: 132
DB: 1 Gaps: 23

US-10-025-380-1081 (1-832) x US-08-188-228-61 (1-2690)

QY      128 GlnSerLysTyrGluGlySerValAlaGlnAsnSerArgPro-GlyLysProPheLeuTy 147
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      441 CAAATGCGAGCCAGAACTCCGCTCGTGTGTCGTTCTTCTGTCAGAGTGTGCTGCT 500
QY      147 rValAsnAlaThrAspLeuAspAspProAlaThrPro----- 159
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      501 AACATCTGGAGAAATTTGAGAC-----TGCATCTCGATTTACACAGAAAGTTTCCA 554
QY      159 ----- 159

```

Db 555 TATCAATCAGCAGCTGTAATTCATTGAGGACCACTCAATTCATACTGACCTTCAGTGA 614
 Qy 160 -----AsnGlyLeuLeuTyrGlnIleValIleGlnLeuProMetIleAsn 176
 Db 615 CTGTAAAGGAAACGACAGCTACCTATGAGGTC-----TC 650
 Qy 176 nValMetTyrPheGlnIleAsnLeuTyrThrGlyAlaIleSerLeu-----ThrAr 193
 Db 651 GAGCCCATCTTCAAGGTGAACAGCATGGCGGCTTGTGCTGTGAGAAACATTAATCTGC 710
 Qy 193 gGluGlySerGlnIleLeuAsnProAlaTyrAsnProSerTyrAsn----- 208
 Db 711 AGTGGGGAACAACTGTGTCCTCATGACAGGACCCCATCGGAGATATGCAAGT 770
 Qy 209 -LeuValIleSerValIleAsnMetGlyGlnSerGlyAsn-----SerPheSerAs 226
 Db 771 CGTATGTCGGGGGAAAGACATCCAGGCTCTCTCGAGATATTAATTAATTTGCAAG 830
 Qy 226 pThrThrSerVal-----AspIleIleValThrGlnAsnIleTyrVal 241
 Db 831 AACTTCCTCTCTCCAAAGACAAAGAGGTCATTTGTGTATCT----- 873
 Qy 241 aProIleProValGlnMetValGlnAsnSerThrAspProIleAspProIleThrG 261
 Db 874 -----CCCATTTTAATTCAGAGATCAGAGACAGCTTCCCAAGATATTTGGCAA 926
 Qy 261 nValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspGlyGlyValProAr 281
 Db 927 GGTA-----GTCATATGTGACAGGCCAGAAAG 953
 Qy 281 gPheProPhe-----SerIleAspGlnGly----- 289
 Db 954 GTCCAGATTCGGGCTCACTGAAAGAGGAGTGAATCAAGGCTTAAGGAATTTTCAGAA 1013
 Qy 290 -----GlyAspIleTyrValThrGlnProLeuAspArgGlnGlyValAspAl 305
 Db 1014 CAATGAGAACACAGGAGCGTCTCCGTGACAGGACCTTGACAGAGATGTAATCGCTGT 1073
 Qy 305 aTyrValPheTyrAlaValAlaTyrAspGlnTyrGlyValProLeuSerTyrProLeu 325
 Db 1074 TTATCAACTATTTGTGAGACCACTGATGTCAATGGCMAAATCTCGAGGGCGGAGGCC 1133
 Qy 325 uIleHisValIleValIleAsnAspAsnProProThrCysProSerProValTh 345
 Db 1134 TCTGAAGATCAATTGATGATCAAGATGACACCGACCATCTTTGGGAGGCCCTTA 1193
 Qy 345 rValPheGlnValGlnGlnLeuAsnGlnTyrGlnGlyAsnSerIleGlyThrLeuThrAla 365
 Db 1194 CATGGGCGACGTCATGGAAGGTCACCCACAGGACCAAGTATGGGATGACAGCTT 1253
 Qy 365 sAspArgAspGlnGlnLeuThrAlaAsnSerPheLeuAsnTyrArgIleValGlnGlnTh 385
 Db 1254 TGATGCGAGATGACCCAGCCGATTAATGCTCTCGGTATATATCCCTCAACAGAC 1313
 Qy 385 rProIleLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAlaGly--MetLeuG 404
 Db 1314 GCTGCAAGACATCTCCCAACATGTTTACATGATCTCTGAGAAAGACATTTGTAC 1373
 Qy 404 nLeuAlaIleGlnSerLeuTyrGlnAspThr-----ProGlnTyrAsnLeuTh 421
 Db 1374 TGTGTGTGACCTGGCTGTGACCGAGAGACTCTGGAATAATCCCAAGTATGAGAT 1433
 Qy 421 rIleGlnValSerAspLys-----AspPheLysThrLeuCysPheValG 436
 Db 1434 CATGAGGCTCAAGATATGCTGATCTGATGTTGATTAACAGGACGCGCACAGCCAC 1493
 Qy 436 nIleAsnValIleAspIleAsnAspGlnIleProIlePheGlnTyrSerAspTyrGlyAs 456
 Db 1494 GATCATGATCATGACAAATATGATCACTCACCAAAATTCACCAAGAAAGGTTT----- 1548
 Qy 456 nLeuThrLeuAlaGlnAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAs 476
 Db 1549 -CAAGCCACAGTCGAGAAAGAGCGTGGGA--CTATTTGTCAATTTGACAGTTGAAGA 1604

Qy 476 pAlaAspGlnProPheThrGlySerSerLeuIleLeuTyrHisIleIleGlyAspAs 496
 Db 1605 TAAGATGACCCACACAGGATGATGAGGAGGCTGCTTACATATATCAACGAAACC 1664
 Qy 496 rGlnGlyArgLeuGlyValAspThrAspProIleThrAsnThrGlyTyrValIleIle 516
 Db 1665 CGGGCAGAGCTTGAATCAATCAACCAACCTCAACCAACCAAGAGGAGTCTTCTGTGT 1724
 Qy 516 sLysProLeuAspPheGlnThrAlaAlaValSerAsnIleValPheLeuAlaGlnLeuP 536
 Db 1725 CAACCATTTGACTATGAATTAATTCCTCCACACCTGCTATCAAACTGGAATAATGA 1784
 Qy 536 oGluProLeuValPheGlyValIleTyrAsnAlaSerSerPheAlaIlePheThrLeu 556
 Db 1785 AGACCCACTGTGACCCGAGCTCTCTTACAGGCCCCAGCTCACAGCCATCCATCAC 1844
 Qy 556 eValThrAspValAsnGlnValaProGlnPheSerGlnHisValPheGlnAlaValSe 576
 Db 1845 TGTCTGAGATGACAGAGGCGCCAGCTTCTTACAGACCCCATGATGATGACAGGCA 1904
 Qy 576 rGluAspValAlaIleGlyThrIleValGlnAsnValThrAlaIleAspProGlnGlyLe 596
 Db 1905 GGAGACCTCTCTGTGGGCGAGCGTGTCTGACAGTGAATGCCAGGACCCGACCTCCT 1964
 Qy 596 uAsp-----IleSerTyrSerLeuArgGlyAspThrArgGlyTyrLeuTyrIleAs 613
 Db 1965 GCAGCATCAACCAATCAGATGATTTCTGTTTCAAGACCCAGCAGGTTGGCTGATATTA 2024
 Qy 613 pHisValThrGlyGlnIlePheSerValAlaProLeuAspArgGlnVala----- 629
 Db 2025 CCCATCAATGAAGGACGTGTGACACACAGCTGTGAGACCGTGAAGTCCCATTTGCGA 2084
 Qy 630 -GlySerProTyrArgValGlnValAlaThrGlnValGlyGlySerSerLeuSerSe 649
 Db 2085 CAACAGGCTGTACACTGCTCTCTCTGCAATTTGACAGTGGCAACCTCCGCTACGGG 2144
 Qy 649 rValSerGlnPheHisLeuIleLeuMetAspValIleAsnAspAsnProProArgLeuAla 669
 Db 2145 CACTGGGACTTCTGTGATTAACCTTGAGAGCGATGAATGACATGACCCCG----- 2193
 Qy 669 sAspTyrThrGlyLeuPhePheCysHisProLeuSerAlaPro----- 683
 Db 2194 -----TTCAATTAACCCACAGTATGAGATGATGATGATGCCAA 2234
 Qy 684 -----GlySerLeuIlePheGlnAlaThrAspAspArgGlnHisLeuPheArgGlyPr 701
 Db 2235 AAACCTCAGTGTGATGATTTTGGAGCGTCAAGTAAAGATCTTACCCGAAATACAGATCC 2294
 Qy 701 oHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAsp---TyrGlnValSerL 720
 Db 2295 T---TTCAATTTGAAATTCACAAACAGCTGTCTCTGATTAAGTCTGGAAGATCTCCA 2351
 Qy 720 sIleAsnGlyThrHisIleArgLeuSerThrArgHisThrAspPheGlnGlnValGAla 740
 Db 2352 GATCAACAAATACACAGCCCTGGTAAGCTTTCTTCAA---AATCTGAACAAACAAACTA 2408
 Qy 740 rValValLeuIleArgIleAsnAspArgGlyAlaGlnProProLeuGlnGlyIleValSerLe 760
 Db 2409 CAACCTGCCATCATGATGTGACAGATTCAGGAAACCCATGACAGATATCACAGATCT 2468
 Qy 760 uProValThrPheCysSerCys 767
 Db 2469 CAGGATCAAGTGTCTCTGTC 2490

RESULT 8

US-08-332-643-55
 : Sequence 55, Application US/08332643

: Patent No. 563634

: GENERAL INFORMATION:

: APPLICANT: Suzuki, Shintaro

: TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS

: NUMBER OF SEQUENCES: 56

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: Two First National Plaza, 20 South Clark
STREET: Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,643
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/872,643
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 5639634 and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/30795
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 2690 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-332-643-55

Alignment Scores:
Pred. No.: 1,596-62 Length: 2690
Score: 660.50 Matches: 205
Percent Similarity: 42.31% Conservative: 103
Best Local Similarity: 28.16% Mismatches: 288
Query Match: 15.29% Indels: 132
DB: 1 Gaps: 23

US-10-025-380-1081 (1-832) x US-08-332-643-55 (1-2690)
QY 128 GlnSerLysTyrGlnGlySerValArgGlnAsnSerArgPro-GlyLysProPheLeu 147
DB 441 CAAATGCAAGCGGAACTCCGCTGTTGTCGCTTCTCTGTCAGGTCGCTGCT 500
QY 147 rValAsnAlaThrAspLeuAspProAlaThrPro- 159
DB 501 AACATGTCAGAGATTGGAC-----TGCACCTCGATTGACAGAGAAAGTTTCA 554
QY 159 ----- 159
DB 555 TATCATCAAGCCAGCGTAATTCATTGAGACGATCAATTCATACTTGACTTGAGTA 614
QY 160 -----AsnGlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAs 176
DB 615 CTGTAAGGAGAAAGACAGCTACGCTATGAGTC-----TC 650
QY 176 nValMetTyrPheGlnIleAsnAsnLysThrGlyAlaIleSerLeu-----ThrAr 193
DB 651 GAAGCCATCTTCAAGGTGAACAGCGATGCGGCTTGTGCTTGAGAAACATACTGCG 710
QY 193 gGlnGlySerGlnGlnLeuAsnProAlaLysAsnProSerTyrAsn----- 208
DB 711 AGTGGGCAAAACTCTGTTGCTTCATGACGAGACCCCATGCGGAGATATGGCAGACT 770
QY 209 -LeuValIleSerValLysAspMetGlyGlnSerGluAsn-----SerPheSerAs 226

```

```

DB 771 CGTATGTCGGGGGAAAGACATCCAGGGCTCTTCGACAGATATTTAAATTTGCAAG 830
QY 226 pThrThrSerVal-----AspIleIleValThrGluAsnLerPrysaI 241
DB 831 AACTCTCTCTGTCGCAAGCAAAAGAGTCATGTTGTAATCT- 873
QY 241 aProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThrG1 261
DB 874 -----CCCATTTTATTTCCAGAAATCAGAACACCTTTCCCAAGATATGTGGCA 926
QY 261 nValArgTyrAsnAspProGlyValaglnTyrSerLeuValAspLysGluLysLeuProAr 281
DB 927 GGTAA-----GTGATAGTACAGCGCCAGAAAG 953
QY 281 gPheProPhe-----SerIleAspGlnGlu----- 289
DB 954 GTCCAAAGTTCCGGCTCACTGGAAGAGAGATGATCAAGCCCTAAAGAAATTTTCAGAA 1013
QY 290 -----GlyAspIleTyrValThrGlnProLeuAspArgGlnGluLysAspAl 305
DB 1014 CAATGAGAACACAGGAGAGCTCTCCGTGACAGGACCTTGACAGAGATATTCGCTGT 1073
QY 305 aTyrValPheTyrAlaValAlaLysAspGlnTyrGlyLysProLeuSerTyrProLeuG1 325
DB 1074 TTAATCAACTATTTGTGGAGACCACTGATGTCATGGCAAAACTCTGAGAGGCGCGGTGCC 1133
QY 325 uIleHisValLysValLysAspIleAsnAspAsnProProThrCysProSerProValTh 345
DB 1134 TCTGGAAGTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
QY 345 rValPheGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 365
DB 1194 CATCGGCCACGTATATGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253
QY 365 sAspArgAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 385
DB 1254 TGATGCAAGATGACCCAGCCACCGATATGCTGCGGTATATATATATATATATATATAT 1313
QY 385 rProLysLeuProMetAspGlyLeuPheLeuIleGlnIleThrTyrAlaGly---MetLeuG1 404
DB 1314 GCCTGACAAAGCCATCTCCCAACATGTTCTATCATGATCTCTGAGAAAGAGATTTGCTAC 1373
QY 404 nLeuAlaLysGlnSerLeuLysGlnAspThr-----ProGlnTyrAsnLeuTh 421
DB 1374 TGTGTGTCACTCGCGCTGTCGACCGAGACTCTGCAAAATCCCAAGATGAACTGAT 1433
QY 421 rIleGlnValSerAspLys-----AspPheLysThrLeuCysPheValG1 436
DB 1434 CATCGAGGCTCAAGATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1493
QY 436 nIleAsnValIleAspIleAsnAspGlnIleProIlePheGlnLysSerAspTyrGlyAs 456
DB 1494 GATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1548
QY 456 nLeuThrIleAlaGluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAs 476
DB 1549 -CAAGCCACAGTCGAGAAAGAGAGCTGTGGA---GTTATTTCAATTTGACACTGTAAGA 1604
QY 476 pAlaSerGlnProPheThrGlySerSerLysIleLeuTyrHisIleIleLysAspSe 496
DB 1605 TAAGATGATACCCACACACAGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1664
QY 496 rGlnGlyArgLeuGlyValAspThrAspProHisThrAsnThrGlyTyrValIleLeu 516
DB 1665 CGGGCAGAGCTTTGAATATCACCAACCTTCAACCAAGAGATCTTCTTCTTGT 1724
QY 516 sLysProLeuAspPheGlnThrAlaValSerAsnIleValPheLysAlaGluAsnPr 536
DB 1725 CAATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1784
QY 536 oGluProLeuValPheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuI1 556

```


Db 1785 AGACCCAGCTGACCCGACCTCTCTACAGGCCCCAGCTCCAGACCAACCTCCACATCAC 1844
 Qy 556 eValThAspValAsnLysLysProGlnPheSerGlnHisValPheGlnLysValSer 576
 Db 1845 TGTCCTGGAGTCAACGAGGCCCCAGTCTTCTACCCAGACCCCATGATGCTGACAGGCA 1904
 Qy 576 rGluAspValAlaIleGlyThrLysValGlyAsnValThrAlaLysAspProGluLys 596
 Db 1905 GGAGAGCTCTCTGTGGGACGCGCTGCTGACAGTGAATGCCAGGACCCGACCTCCCT 1964
 Qy 596 uAsp-----IleSerTySerLeuArgGlyAspThrArgGlyTyrLeuLysIleAs 613
 Db 1965 GCAGCATCAACACCATCAGTATCTGTTTCAAGAGACCCAGAGGTGGCTGATATATAA 2024
 Qy 613 pHisValThrGlyGluIlePheSerValAlaProLeuAspArgGluLys----- 629
 Db 2025 CCCCATCAATGGGACTGTTGACACACACAGCTGCTGACCGTGAAGTCCCATTTGTGCA 2084
 Qy 630 -GlySerProTyrArgValGlnValAlaThrGlnValGlyGlySerSerLeuSerSe 649
 Db 2085 CAACAGCGTGAACCTGCTCTCTCTCTGCGCAATTGACAGTGGCAACCTCCGCTAGCGG 2144
 Qy 649 rValSerGluPheHisLeuIleLeuMetAspValAsnAspAspProArgLeuAlaLys 669
 Db 2145 CACTGGAGCTTGTGCTGATTAACCTCGAGAGAGTGAATGACATGCCCGC----- 2193
 Qy 669 sAspTyrThrGlyLeuPhePheCysHisProLeuSerAlaPro----- 683
 Db 2194 -----TTCAATTATCCCAAGTGAAGTGAAGTGTGATGATGATGCA 2234
 Qy 684 -----GlySerLeuIlePheGlnAlaThrAspAspAspGlnHisLeuPheArgGlyPr 701
 Db 2235 AAACCTCAGTGTACTATTGGAGCATCAGATGAGATTTTCAACCCGAAATACAGATCC 2294
 Qy 701 oHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAsp--TyrGluValSerLys 720
 Db 2295 T--TTCAATTGAAATCCCAACAAAGCTGTCTGATGAAGTCTGGAAGATCTCCAA 2351
 Qy 720 gIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGlnAlaGlyLys 740
 Db 2352 GATCAACAATGACACGCCCTGTAGAGCTTTCTTCAA--AATGTGAACAAAGCAAACTA 2408
 Qy 740 rValValLeuIleArgIleAsnAspGlyGlyArgProLeuGluLysIleValSerLe 760
 Db 2409 CAACCTGCCATCATGTGTGACAGATTGACGGAACCAACCATGACGAATATCAAGATCT 2468
 Qy 760 uProValThrPheCysSerCys 767
 Db 2469 CAGGCTACAAGTGTCTCTGC 2490
 RESULT 9
 US-08-332-638-61
 ; Sequence 61, Application US/08332638
 ; Patent No. 5646250
 ; GENERAL INFORMATION:
 ; APPLICANT: Suzuki, Shintaro
 ; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
 ; NUMBER OF SEQUENCES: 62
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ; ADDRESSEE: Borun
 ; STREET: 6300 Sears Tower, 233 S. Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/332,638

; FILING DATE: 01-NOV-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/872,643
 ; FILING DATE: 17 APR 1992
 ; APPLICATION NUMBER: US/08/049,460
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 5646250and, Greta E.
 ; REGISTRATION NUMBER: 35,302
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (312) 474-6300
 ; TELEFAX: (312) 474-0448
 ; TELEEX: 25-3856
 ; INFORMATION FOR SEQ. ID NO: 61:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2690 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-08-332-638-61
 Alignment Scores:
 Pred. No.: 1,596-62 Length: 2690
 Score: 660.50 Matches: 205
 Percent Similarity: 42.31% Conservative: 103
 Best Local Similarity: 28.16% Mismatches: 288
 Query Match: 15.29% Indels: 132
 DB: 1 Gaps: 23
 US-10-025-380-1081 (1-832) x US-08-332-638-61 (1-2690)
 Qy 128 GlnSerLysTyrGluGlySerValAlaGlnAsnSerArgPro-GlyLysPheProheLeuTyr 147
 Db 441 CAATAAGCAAGCCGAGAACTCCGCTCGTCTGTGCGCTTCTCTCCAGGTGCTGCTGCT 500
 Qy 147 rValAsnAlaThrAspLeuAspAspProAlaThrPro----- 159
 Db 501 AACATCTGCAGAAATTTGGAC-----TGCATCTCTGATTTACAGCAAAAGTGTCCA 554
 Qy 159 ----- 159
 Db 555 TATCATACAGCCAGCGAATTGATGAGACAGTCAATTCTAACTTGACCTTGATGA 614
 Qy 160 -----AsnGlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAs 176
 Db 615 CTGTAAGGGAACGACAGCAAGCTACGCTATGAGTGC-----TC 650
 Qy 176 nValMetTyrPheGlnIleAsnAsnLysThrGlyValAlaIleSerLeu-----ThrAr 193
 Db 651 GAGCCCATACTTAAGGTGAACAGCATGCGGCTTAGTGTCTCTGAGAAACATAACTGTC 710
 Qy 193 gGluGlySerGlnGluLeuAsnProAlaLysAsnProSerTyrAsn----- 208
 Db 711 AGTGGGCAAAACTCTGTGTCATGACAGACCCCATGGGGAAGATATGCGAGACT 770
 Qy 209 -LeuValIleSerValLysAspMetGlyGlnSerGluAsn-----SerPheSerAs 226
 Db 771 CGTGAATGTCGGGGGGAAGACATCCAGGCTCTCTTCCAGGATATTTAAATTTGCAAG 830
 Qy 226 pThrThrSerVal-----AspIleIleValThrGluAsnIleTyrLysAl 241
 Db 831 AACCTCTCTGCTCCCAAGACAAAGAGTCCATTTGTGATCT----- 873
 Qy 241 aProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThrGln 261
 Db 874 -----CCCATTTTAATTCAGAGATCAGACAGCGCTTTCCCAAGAGATGTGGCAA 926
 Qy 261 nValArgTyrPheAsnAspProGlyValGlnTyrSerLeuValAspLysGluLysLeuProAr 281
 Db 927 GGTG-----GTCGATGATGACAGGCGCCAGAAAG 953

QY 281 gphepPrope-----SerileaspGlnGlu-----289
 Db 954 GTCCAGTTCGCGCTCACTGAAAGGAGTGATCAAGAGCTTAAGAAATTTTTCAGAT 1013
 QY 290 -----GlyaspIleTyValThrGlnProLeuAspArgGluGlyAspAl 305
 Db 1014 CAATGAGAACACAGAGGCGCTCCGTGACACGACCTTGGACAGAAAGTAATCGCTGT 1073
 QY 305 aTyValPheTyValAlaValAlaValAspGluTyGlyLeuProLeuSerTyProLeuG 325
 Db 1074 TTATCAACTATTTTGGAGACCACTGATGTCAATGGAAACTCTCCAGGGGCGGTGCC 1133
 QY 325 uileHleValIleValIleAspIleAspAspAspProProCysProSerProValTh 345
 Db 1134 TCTGAGAGTCACTGTGTGATGATCAAGATGACACGACCATCTTTGGGAGGCCCTTA 1193
 QY 345 rValPheGluValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 365
 Db 1194 CATGGCCACGTCATGGAGGTCACCCACAGCCACAGTCATGCGAGTGAACAGCTT 1253
 QY 365 sAspArgAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 385
 Db 1254 TGATGCAATGACCCACCCACCATATGCTCTCGGTATATATCCGTCAACAGAC 1313
 QY 385 rProTyLeuProMetAspGlyLeuPheLeuIleGlnThrTyAlaGly--MetLeuG 404
 Db 1314 GCCGCAAGCCATCTCCCAACATGTTCTCATGATGATCCGAGAAAGAGACATGTGCAC 1373
 QY 404 nLeuAlaValGlnSerLeuTyGlnAspThr-----ProGlnTyAsnLeuTh 421
 Db 1374 TGTGTGTCACTGCGCTGGAGCCAGAGACTCTGAAATATCCCAATATGAAATCGAT 1433
 QY 421 rIleGluValSerAspLys-----AspPheLysThrLeuCysPheValG 436
 Db 1434 CATGAGGCTCAAGATATGCTGAGCTGATGTTGATTAACAGGACCGGACACAGCCAC 1493
 QY 436 nIleAsnValIleAspIleAsnAspGlnIleProIlePheGlnTySerAspTyGlyAs 456
 Db 1494 GATCATGATCGATGACAAAATATGATCACTACCAAAATTCACCAAGAAAGATTT----- 1548
 QY 456 nLeuThrLeuAlaGlnAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAs 476
 Db 1549 -CAAGCCACAGTCGAGGAGAGAGCTGGGA--GTATTTGTCATTAATGACAGTTGAAGA 1604
 QY 476 rAlaAspGluProPheThrGlySerSerTyIleLeuTyHisIleIleTyGlyAspSe 496
 Db 1605 TAAAGATGACCCCAACAGAGTGATGAGAGGCTGCTTACACCATTCATCAACGAAACCC 1664
 QY 496 rGlnGlyArgLeuGlyValAspThrAspProHisIleThrAsnThrGlyTyValIleIle 516
 Db 1665 CGGGCAGAGCTTGAATCAATCAACACCACTCAAAACCAAGAGAGATGCTTCTGTGT 1724
 QY 516 sTyAspProLeuAspPheGlnThrAlaAlaValSerAsnIleValPheTyAlaGluAsp 536
 Db 1725 CAATCACTGATGATGATAATTTCTGCTTCCACACCTGCTGATCAAGTGGAAATGA 1784
 QY 536 ogIuProLeuValPheGlyValIleTyAsnAlaSerSerPheAlaLysPheThrLeuI 556
 Db 1785 AGACCCCACTGTCACCGACGCTCTCAACGCGCCAGCTTCAACAGCCACCTTCAACATCAC 1844
 QY 556 eValThrAspValAsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSe 576
 Db 1845 TGTCTGAGATGTCAACAGGGCCCACTTCTTACCCAGACCCCATGATGTGACAGGCA 1904
 QY 576 rGluAspValAlaIleGlyThrIleValGlyAsnValThrAlaLysAspProGluGlyLe 596
 Db 1905 GGAGGACCTCTCTGCTGGCAGCGGTGCTGACAGTGAATGACACGAGCCCACTCCCT 1964
 QY 596 uAsp-----IleSerTySerLeuArgGlyAspThrArgGlyThrLeuLysIleAs 613
 Db 1965 GCAGCATCAAAACCATCAAGGATATCTGTTTCAAGAGACCCAGAGGTGGTGGTGAATTA 2024

QY 613 pHsValThrGlyGlnIlePheSerValAlaProLeuAspArgGluAla----- 629
 Db 2025 CCCCATCAATGGAGACTGTGTACACACAGCTGTGCTGACCGTGAAGTCCCATTTTGA 2084
 QY 630 -GlySerProTyArgValGlnValAlaIleThrGluValGlyGlySerSerLeuSerSe 649
 Db 2085 CAACAGGCTGTACACGCTCTCTCCGCAATGTACAGTGGCAACCTCCGCTACGGG 2144
 QY 649 ValSerGluPheHisLeuIleLeuMetAspValAspAspAspProProArgLeuAla 669
 Db 2145 CACTGGAGACTTGTGTATACCTTGGAGACGTAATGACAAATGATCCCG----- 2193
 QY 669 sAspTyThrGlyLeuPhePheCysHisProLeuSerAlaPro----- 683
 Db 2194 -----TTCAATTAACCCACAGTACGTGAAGTCTGTATGATGCCA 2234
 QY 684 -----GlySerLeuIlePheGlnAlaThrAspAspGlnHisLeuPheArgGlyPr 701
 Db 2235 AAACCTCAGTGTAGTATTTTGGAGCATGATGATTAAGATCTTCAACCGAATACAGATCC 2294
 QY 701 ohIlePheThrPheSerLeuGlySerGlySerLeuGlnAspAsp---TrpGluValSerTy 720
 Db 2295 T---TTCAATTTGAAATTCACAAACAGCTGTCTCGATTAAGTCTGGAAGATCTCAA 2351
 QY 720 sIleAsnGlyThrHisAlaArgLeuSerThrArgHisIleThrAspPheGlnGluAlaTy 740
 Db 2352 GATCAACAATACACAGCCCTGTGATGCTTCA--AACTGAACAACAAACTA 2408
 QY 740 rValValLeuIleArgIleAsnAspGlyArgProProLeuGlnGlyIleValSerIle 760
 Db 2409 CAACCTGCCCATCATGATGTGACAGATTCAGGGAACACACCATGATGATATTCACAGATCT 2468
 QY 760 uProValThrPheCysSerCys 767
 Db 2469 CAGGTTCAAGTGTGCTCTGC 2490

RESULT 10
 US-08-474-067-3
 ; Sequence 3, Application US/08474067
 ; Patent No. 5811518
 ; GENERAL INFORMATION:
 ; APPLICANT: Ranscht, Barbara
 ; TITLE OF INVENTION: T-Cadherin Adhesion Molecule
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Campbell and Flores
 ; STREET: 4370 La Jolla Village Drive, Suite 700
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: United States
 ; ZIP: 92122
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/474,067
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/213,361
 ; FILING DATE: 14-MAY-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/607,293
 ; FILING DATE: 30-OCT-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Campbell, Cathryn A.
 ; REGISTRATION NUMBER: 31,815
 ; REFERENCE/DOCKET NUMBER: P-LJ 1682
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 535-9001
 ; TELEFAX: (619) 535-8949

DB 1888 CCA---TTTAATTTGAAGTGAAGCAATCTGTCAGAAAGATTAGGAATCAAC 1944
 QY 720 LysleleenglyThrHisAlaArgLeuSerThrArgHisThrAspPheGluAla 739
 DB 1945 AAGCTTAAACATCTACTATGCCCCAGGTGCTGCTTCAA--AACCTGAAAAAGCCCAAT 2001
 QY 740 TyrValValleuLeuArgLeuAsnAspGlyGlyArgProPheGluGlyIleValSer 759
 DB 2002 TACAACATCCCAATCTCAGTACAGATTCTGGAAACACCTCTGACTAACACACGAA 2061
 QY 760 LeuProValThrPheCysSerCys 767
 DB 2062 CTGAATTACAGATGCTGCTCTGC 2085

RESULT 11

US-08-474--068A-3
 Sequence 3, Application US/08474068A
 Patent No. 5837525
 GENERAL INFORMATION:
 APPLICANT: Ranach, Barbara
 TITLE OF INVENTION: T-Cadherin Adhesion Molecule
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Campbell & Flores LLP
 STREET: 4370 La Jolla Village Drive, Suite 700
 CITY: San Diego
 STATE: California
 COUNTRY: United States
 ZIP: 92122
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/474,068A
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/213,361
 FILING DATE: 14-MAY-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/607,293
 FILING DATE: 30-OCT-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P-LJ 1683
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2779 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 40..2191
 US-08-474-068A-3

Alignment Scores:

Pred. No.: 3,44e-59 Length: 2779
 Score: 630.50 Matches: 193
 Percent Similarity: 44.77% Conservative: 115
 Best Local Similarity: 28.05% Mismatches: 283
 Query Match: 14.59% Indels: 97
 DB: 2 Gaps: 25

US-10-025-380-1081 (1-832) x US-08-474-068A-3 (1-2779)

QY 150 AlaThrAspLeuAspAspProAlaThrProAsnGlyGln-----LeuTyrTyrGlnIle 167
 DB 103 GCGAAGATTTAGAA-----TGACCCCTCGAATTCACGAAAAAGTTTATN----- 150
 QY 168 ValIleGlnLeuProMetIleAsnValMetTyrPheGlnIleAsnValSerThrGly 187
 DB 151 ---ATTGAACAGCA-----TTGAAATTCACAGAGACACAGCA 186
 QY 188 AlaIleSerLeuThrArgGlu-----GlySerGlnIleuAsn-----ProAlaLys 203
 DB 187 ATTCTGAACCTGGTGTGTTGATGACTGCAGAGGAATTAACAATTGAATCTTGAAATTTCT 246
 QY 204 AsnProSerTyrAsnLeu-----ValIleSerValLysAspMetGly--- 217
 DB 247 AACCCAGACTTAAAGTGAACAGATGATCTTTAGTTCACATGAAGATGATACGA 306
 QY 218 -----GlyGlnSerGluAsnSerPheSerAspThrThrSer 229
 DB 307 GCTGCAGAGCTTGTGTTGTCATGACGCGTCTGAGCATGCT---GAGATATGCGAGAA 363
 QY 230 ValAspIleIleValThrGluAsnIleThrLysAlaProLys----- 243
 DB 364 ATTTGATTTGTGAGCTGATGAGAAAGACAGATCATTAAGAAATCTTTAAGATGAA 423
 QY 244 -----ProValGluMetVal 248
 DB 424 GGCACCTTGGAATTCACAAACAAAAAGGCTAATCTGGCGACTCCAAATTAATTCGA 483
 QY 249 GluAsnSerThrAspProHisProIleLysIleThrGlnValArgTrpAsnAsp----- 266
 DB 484 GAAATCAAAAGACACCATTTCCAGATCAGTTGGCAAGGTCAACAGAGTGAAGGACA 543
 QY 267 ProGlyAlaGlnTyrSerLeuValAspLys-----GluLysLeuProAlaGlyPhePro 284
 DB 544 GAGGGAGCAAAAGTCCGACTCTGTTAAGAGATGATGAACCCGAAAGAAATTTT 603
 QY 285 SerIleAspGln---GluGlyAspIleTyrValThrGlnProLeuAspArgGluGlnLys 303
 DB 604 AGAATCAATGATGATGAGTGGAGTGTCTGTGATCCGACCCCTGGATTAAGAACATGA 663
 QY 304 AspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeuSerTyrPro 323
 DB 664 GCCAATTAATGAGCTGGAAGTGAAGTAAGCATTAAGTGAAGGAAATCATTAATGAGCCA 723
 QY 324 LeuGlnIleHisValLysValIleAspIleAsnAspAsnProThrCysProSerPro 343
 DB 724 GTCCGCTGATATTTCTGTATTGATCAAAATGAATACAGGCGGATGTTCAAGAAAGA 783
 QY 344 ValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGlyThrLeuThr 363
 DB 784 CCTATGTTGTCATGATGATGAGATGATCCCTACAGAAACAATGATGCGGATGACA 843
 QY 364 AlaHisAspArgAspGluGlnAsnThrAlaAsnSerPheLeuAsnTyrArgIleValGlu 383
 DB 844 GCATTTGATGCTGATGATCTGACACAGCAAGCTCTTCTGCGGATTAACATCCTCAAG 903
 QY 384 GlnThrProLysLeuProMetAspGlyLeuPheLeuIleGln----- 397
 DB 904 CAGACACCTACCAAACTTCCCAATATGTTCTTACATTGACCCAGAAAAAGGAGATATT 963
 QY 398 ---ThrTyrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrPro 416
 DB 964 GTCACAGTGTGTCGCTGCTGATCTGATCGTGAATCGTGAACATG-----GAAAGCGCG 1014
 QY 417 GlnTyrAsnLeuThrIleGluValSerAspLys-----AspPheLysThr 431
 DB 1015 AAGTACAGAGCTGTTATTGAGCAAGATATGCGGCGCATGATGAGGACTTAATCGGA 1074
 QY 432 LeuCysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLys 451
 DB 1075 ACTGCACTGCCACTTCTTATTGATGACAAAAAGACCCACCCAGATTTACCAAG 1134
 QY 452 SerAspTyrGlyAsnLeuThrLeuAlaGluAspThrAsnIleGlySerThrIleLeuThr 471

```

Db      1135 AAGGAGTT---CAGGCCACAGTAAAGAA-----GAGTCCACAGAGATTAATGTAAAC 1185
Qy      472  TleGlnAlaThrAspAlaAspGluProPheThrGlySerSerIleLeuTyRHisIle 491
Db      1186 TTAACTCTTGATGACCCAGATGACCCAGCAATGAGCATGAGCATGAGCTTCACTATT 1245
Qy      492  TleGlyGlyAspSerGluGlyArgLeuGlyValAlaSerPheAspProHisThrAspThrGly 511
Db      1246 ATTAAACGAAATCCAGGCGCAGAGTTTGGAAATCCATCCATCCCGAGACTTAAGAGGA 1305
Qy      512  TyrValIleIleIleGlySerProLeuAspPheGluThrAlaAlaValSerAsnIleValPhe 531
Db      1306 ATGCTCTCTGTGTCGCAACCTTTAGACTATGAGATTTCAGAGATTTCACACTTCTGATA 1365
Qy      532  LysAlaGluAsnProGluProLeuValPheGlyValIleGlyTyrAsnAlaSerSerPheAla 551
Db      1366 AAAGTAGAAATAGAAAGCCCGTATTCCAGACATAGCTAGCTCCAGTTCCACAGCA 1425
Qy      552  LysPheThrIleuIleValThrAspValAsnGluAlaProGlnPheSerGlnHisValPhe 571
Db      1426 ACAGTTCAGATCACCGTTGAGAGATGATGATGAGAGCCCTGTTTCCACCCAAACCCATG 1485
Qy      572  GlnAlaLysValSerGluAspValAlaIleGlyThrIleValGlyAsnValThrAlaLys 591
Db      1486 ACAGTGAACAAACAGAGAGAACATCCCTATTGGCAGCATTTGTTAAACAGTAAATGCCACT 1545
Qy      592  AspProGluGlyLeuAsp-----IleSerTyrSerLeuArgGlyAspThrArgGly 608
Db      1546 GATCCAGATCTTCTTCAACATCAGATCAGATCAGTATTCAGTTTCAAGAGATCCAGCAAGC 1605
Qy      609  ThrLeuValIleAspHisValThrGlyGluIlePheSerValAlaProLeuAspArgGlu 628
Db      1606 TGGCTAGAGATTAATCCACCAATGATGATCCCTGCCACCATCTGCTGCTGATCGGGA 1665
Qy      629  Ala-----GlySerProTyrArgValGlnValAlaThrGluValGlyGly 644
Db      1666 TCTCCGATGTTCCAGATTAACAATACACTGCTCTCTCCGCAATAGACAGTGGTAAAC 1725
Qy      645  SerSerLeuSerSerValSerGluPheIleLeuIleLeuMetAspValAsnAspAsnPro 664
Db      1726 CTTCTCTCTACAGTACAGAACTTTACATCACTTTCAGTGAAGCTCAATGACATGCTC 1785
Qy      665  -----ProAlaGluAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
Db      1786 CCTCCCTTTACCAACAACCTGGCAAA-----GCTGTGATGATGCT 1827
Qy      681  SerAlaProGlySerLeuIlePheGluAlaThrAspAspArgGlnHisLeuPheArgGly 700
Db      1828 AAAGATCTCAGAGTAGTGGTTCTAGGAGCATCAGCAAAAGCTCCATCCCAACACAGAT 1887
Qy      701  ProHisPheThrPheSerLeuGlySerIleLeuGlnAsnAspTyrGluValSer 719
Db      1888 CCA---TTTAAATTTGAACTGAGTAGCAATCTGCTCCAAAGATTAGAGATCAAC 1944
Qy      720  LysIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluAlaArgAla 739
Db      1945 AAGCTTAACAATACATCTATGCCCAGAGTGTCTGCTTCAA---AACCTGAAGAAAGGCCAAT 2001
Qy      740  TyrValValIleuIleArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSer 759
Db      2002 TACAAACATCCCAATCTCAGTAGCAGATTTCTGAAACACACTCTGACATAACACAGAA 2061
Qy      760  LeuProValThrPheCysSerCys 767
Db      2062 CTGAATTAACAAGTGTCTCTGC 2085

```

RESULT 12
 US-08-472-481-3
 ; Sequence 3, Application US/08472481
 ; Patent No. 5863804
 ; GENERAL INFORMATION:
 ; APPLICANT: Ranscht, Barbara

```

; TITLE OF INVENTION: T-Cadherin Adhesion Molecule
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,481
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/213,361
; FILING DATE: 14-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/607,293
; FILING DATE: 30-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2779 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 40..2191
; US-08-472-481-3
;
; Alignment Scores:
; Pred. No.: 3,44e-59 Length: 2779
; Score: 630.50 Matches: 193
; Percent Similarity: 44.77% Conservative: 115
; Best Local Similarity: 28.05% Mismatches: 283
; Query Match: 14.59% Indels: 97
; Gaps: 25
;
; US-10-025-380-1081 (1-832) x US-08-472-481-3 (1-2779).
Qy      150  AlaThrAspLeuAspAspProAlaThrProAsnGlyGln-----LeuTyRtyrGlnIle 167
Db      103  GCAGAAAGATTAGAA-----TCACCCCTCGAATTCACGAAAGATTATTAT----- 150
Qy      168  ValIleGlnLeuProMetIleAsnAsnValMetTyrPheGlnIleAsnAsnLysThrGly 187
Db      151  ---ATTGAACAGCA-----TTTCAATTCCACAGAGACACCA 186
Qy      188  AlaIleSerLeuThrArgIle-----GlySerGlnIleuAsn-----ProAlaLys 203
Db      187  ATTCTGAACCTGTTGTTGATGACTGCAAGGGAATTAACAATTGAATTCGAAGTTTCT 246
Qy      204  AsnProSerTyrAsnLeu-----ValIleSerValIysAspMetGly--- 217
Db      247  AACCCAGACTTAAGTGAAGCAACAGATGATCTTTAGTTGCTACGTGAAGATGTATCAGAA 306
Qy      218  -----GlyGlnSerGluAsnSerPheSerAspThrThrSer 229
Db      307  GCTGGCAGACCTTTGTTGTCATGCACGCTGTGAGCATCT---GAGGATATGACAGAA 363
Qy      230  ValAspIleIleValThrGluAsnIleTyrIleAlaProLys----- 243

```

Db 364 ATTTGATGTGTGGAGCTGATGAGAGACAGATGCAATTAAGAAATCTTTAAGATGAA 423
244 -----ProValIGluMetVal 248
Db 424 GGCAACCTTGGAATTCACAGACAAAAAGGCTATTCTGGGAGCTCCAAATTTAATTTCA 483
249 GluAsnSerThrAspProHisProIleValIleThrGlnValArgTrpAsnAsp----- 266
484 GAAATCAAGACACCATTTCCAGATCAAGTTCAGTGGCAATCATCAGAGTGAAGGACA 543
267 ProGlyAlaGlnIleThrSerLeuValAspLys-----GluLysLeuProAlaGlnProPhe 284
544 GAGGAGACAAAGTTCACACTCTCTGTAGGAGTACATCAAGACCCGAAAGAAATTTT 603
285 SerIleAspGln-----GluGlyAspIleIleThrValThrGlnProLeuAspArgGluLys 303
604 AGAATCATAGATGATCAAGTGGAGATGCTCTGTGACCCGACCCCTGTGATGAGAAACATA 663
304 AspAlaIleThrValPheThrValAlaValAlaLysAspGluIleThrGlyLysProLeuSerTrpPro 323
664 GCCAATATTAGAGCGGAAAGTTGAAGTAAAGCATTTAAGTGGAAATCATTTGATGGCCA 723
324 LeuGluIleHisValLysValLysAspIleHisAspAsnProProThrCysProSerPro 343
724 GTCCGCTAGATATTTCTGTATTGATCAAAATGATTAACAGCCGAGTTCAAAGAAAGA 783
344 ValIleThrValPheGluValGlnGluAsnGluArgLeuLysAsnSerIleGlyThrLeuThr 363
784 CCTATGTTGTGTCACGTACGATGAGAAAGATCCCTTACAGAACATCTGTGATGCGGATACA 843
364 AlaHisAspArgAspGluAsnThrAlaAsnSerPheLeuAsnTrpArgIleValGlu 383
844 GCATTTGATGCTGATGATCTCTAGACACAGACAGCTCTTCTGCGGTATTAATCTCTCAG 903
384 GlnThrProLysLeuProMetAspGlyLeuPheLeuIleGln----- 397
904 CAGACACCTACCAAACTCCCAAAATATGTTCTTACATTCACCCAGAAAGGAGATATT 963
398 ---ThrThrAlaGluMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrPro 416
964 GTCCAGATGCTGCTGCTGATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1014
417 GlnThrAsnLeuThrIleGluValSerAspLys-----AspPheLysThr 431
1015 AAGTACAGAGCTGTTATTGAAGCCAAAGATATGGGCGGCAATGATGCGAATTTACTGCA 1074
432 LeuCysPheValGlnIleAsnValIleAspIleHisAspGlnIleProIlePheGluLys 451
1075 ACTGCAACCTGCACTATTCTTATGATGATCAAAAACAGACACCCAGACAGATTTACCAAG 1134
452 SerAspIleThrGlnLeuThrLeuAlaGluAspThrAsnIleGlySerThrIleLeuThr 471
1135 AAGGAGTTT---CAGGCCACAGTAAAGAA---GAACTCACAGAGTAAATTAAC 1185
472 IleGlnAlaIleThrAspAlaAspGluProPheThrGlySerSerLysIleLeuThrHisIle 491
1186 TTAACTGTGTGATCCAGATGACCCGCACTGACAGACATGAGAGCTGTCTTACACTATT 1245
492 IleLeuGlyAspSerGluGlyArgLeuGlyValAspThrAspProHisThrAsnThrGly 511
1246 ATTAAACGGAATCCAGGAGAGATTTGAATATCATATCAATCCCACTGCAATTAAGAGGA 1305
512 TyrValIleIleLysLysProLeuAspPheGluThrAlaAlaValSerAsnIleValPhe 531
1306 ATGCTCTCTGTGTCAAACTTTAGATTAAGATTTACACATTTTCAACATTTGCTGATA 1365
532 LysAlaGluAsnProGluProLeuValPheGlyValLysTrpAsnAlaSerSerPheAla 551
1366 AAAGTGAAGAAATGAAGCCGTTGATTCAGACATACCTACCGTCCCAAGTTCCACAGCA 1425
552 LysPheThrLeuIleValIleThrAspValAsnGluAlaProGlnPheSerGlnHisValPhe 571

Db 1426 ACAGTTCAGATCACCTGTGAGATGATGAATGAAGGCCCTGTTTCCACCCAAACCAANG 1485
572 GlnAlaLysValSerGluAspValAlaIleGlyThrLysValGlyAsnValThrAlaLys 591
1486 ACAGTGAACAAACAAAGAAACATCTCTATTGGAGACATTTGTTAACAGTAATGCCACT 1545
592 AspProGluGlyLeuAsp-----IleSerTrpSerLeuArgGlyAspThrArgGly 608
1546 GATCCAGATATCTTTGCAACATCAGATCAGATCAGATATTCAGTTTCAAGATCCAGCAAGC 1605
609 TrpLeuLysIleAspHisValThrGlyGluIlePheSerValAlaProLeuAspArgGlu 628
1606 TGCTAGAGATTAATCCCAACAAATGATCCGTTCCACCACTGCTGCTGATTCGGGAA 1665
629 Ala-----GlySerProTrpArgValGlnValAlaIleThrGlnValGlyGly 644
1666 TCTCCGATGTTGAGATTAACAAATACACTGCTCTCTCTGCAATAGACAGTGTAC 1725
645 SerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspValAsnAspAsnPro 664
1726 CTTCTGTGATCAGATGATCAGAACTTTACATCATCATCCTTGAGAGACGTCATGACATGTC 1785
665 -----ProArgLeuAlaLysAspTrpThrGlyLeuPhePheCysHisProLeu 680
1786 CCTCCCTTACCCCAACACTGCGCAAA-----GTCGTGTGATGATGCT 1827
681 SerAlaProGlySerLeuIlePheGluAlaIleThrAspAspAspGlnHisLeuPheArgGly 700
1828 AAGATCTCAGATGATGTTCTTACGAGACATTCAGACAAAGATCTCCATCCCAACACAGAT 1887
701 ProHisPheThrPheSerLeuGlySerGlySer---LeuGlnAsnAspTrpGluValSer 719
1888 CCA---TTTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1944
720 LysIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluAlaGly 739
1945 AAGCTTAAACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2001
740 TyrValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSer 759
2002 TACAACTCCCAATTCAGAGACAGATTTCTGGAACCACTGTGACTTAACACACAGAA 2061
760 LeuProValThrPheCysSerCys 767
2062 CTGAATTTCAAGTGTGCTCTGC 2085

RESULT 13
US-08-474-067-1
; Sequence 1, Application US/08474067
; Patent No. 581518
; GENERAL INFORMATION:
; APPLICANT: Ranscht, Barbara
; TITLE OF INVENTION: T-Cadherin Adhesion Molecule
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,067
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/213,361
; FILING DATE: 14-MAY-1994

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/607,293
 FILING DATE: 30-OCT-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P-LJ 1682
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3959 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 45..2181
 US-08-474-067-1

Alignment Scores:
 Pred. No.: 6,22e-59 Length: 3959
 Score: 630.50 Matches: 193
 Percent Similarity: 44.77% Conservative: 115
 Best Local Similarity: 28.05% Mismatches: 283
 Query Match: 14.59% Indels: 97
 Gaps: 25

US-10-025-380-1081 (1-832) x US-08-474-067-1 (1-3959)

QY 150 Alathraspleuapaprolath:Proasnglygin-----LeuTYRtyrGlnIle 167
 DB 108 GCAGAAAGTTTGA-----TGCACCCCTGGATTCCAGCAAAAGGTTTITAT----- 155
 QY 168 ValIleGlnleuPromeTlleasnaValMetTYRpheGlnIleasnaValtyrGly 187
 DB 156 ---ATTGAACAGCA-----TTTGAATTCACAGAGACAGCA 191
 QY 188 AlalIleSerleuThraGlu-----GlySerGlnIleuAsn---ProAlIlys 203
 DB 192 ATTCTGAACCTGGTGTGGATGACTGCAGAGGGAATTAACAATTGAACCTTGAAGTTTCT 251
 QY 204 AspProSerTYraAsnleu-----ValIleSerValIysAspMetGly--- 217
 DB 252 AACCCAGACTTAAAGTGAACACAGATGATCTTAACTGACATGAAGATATCAAA 311
 QY 218 -----GlyGlnSerGluAsnSerPheSerAspThrThrSer 229
 DB 312 GCTGGCAGAGCTTGTGTTCATGCACCGCTGAGCATGCT---GAGGATATGCGA 368
 QY 230 ValAspIleIleValIThrGluAsnIleTrrIyAlaIAspIlys----- 243
 DB 369 ATTTTGATTGTGGAGCTGATGAGAGACGATGATTAAGAAATCTTTAAGATGAA 428
 QY 244 -----ProvalGluMetVal 248
 DB 429 GGCACCTTGGAAATTCAGACAAAAGGCTATTCTGGCGACTCCAAATTTATTTCA 488
 QY 249 GluAsnSerThraPProHAspProIleIySleIleThrGlnValIAspTrpAsp----- 266
 DB 489 GAAATATGAAGACACACATTTCCAGATCACTTGGCAAGTCAACAGAGTGAAGGAGCA 548
 QY 267 ProGlyAlaGlnTYrSerleuValAspIlys-----GluIySleuProAspPheProPhe 284
 DB 549 GAGGAGACCAAGTTCGCACTCTCTGATGAGAGATGATCAAGACCCGAAAGGAAATTTT 608
 QY 285 SerIleAspGln---GluGlyAspIleIyTYrValIThrGlnProIleuAspAspGluGlnIlys 303
 DB 609 AGAATTCATGATGATCACTGGGAGTGTCTCTGTGACCCGACCTCGAGATGAGAAAGCAATA 668
 QY 304 AspAlaTYrValIlePheTYrAlaValAlaIyAspGlnTYrGlyIySleProIleSerTYrPro 323

DB 669 GCCAATTATGAGCTGGAGAGTTGAAGTAACGGAATTTAAGTGGAAATATCATGTATGCGCCA 728
 QY 324 LeuGlnIleIleValIyValIyAspIleAsnaAspAspProThrCysProSerPro 343
 DB 729 GTCCGCTTGAATATTTCTGTATTGATCAAAATGATTAACAGCCGATTTCAAAAGAGA 788
 QY 344 ValIThrValIlePheGluValIleGlnIleuAsnIyAspSerIleGlyIThrLeuThr 363
 DB 789 CCTATGTTGTGCAGCTCATGAGAGATCCCTCAAGAAACAACGTGTATGCGATGACA 848
 QY 364 AlahIAspAspAspGlnIleuAsnIleAsnSerPheLeuAsnTYrArgIleValIleGlu 383
 DB 849 GCATTTGATGCTGATGATCTTACACAGCAACGCTTCTCGGATTAACATCTCAAG 908
 QY 384 GlnThrProIySleuPromeTleAspGlyLeuPheIleGln----- 397
 DB 909 CAGACACCTTACCAAACTTCCCAATATGTTCTTACATTTGACCCAGAAAGGAGATATT 968
 QY 398 ---ThrTYrAlaGlyMetLeuGlnIleuAlaIyGlnSerIleuIySleGlnAspThrPro 416
 DB 969 GTCAAGTGTGTCACTGTATCTGTATGATGAGATGAGATG-----GAAACGCGC 1019
 QY 417 GlnTYraAsnleuThrIleGlnValIleAspIlys-----AspPheIyThr 431
 DB 1020 AAGTACGAGCTGGTATTGAAGCCAGATATGCGCGCATATGTGGACTTACTGCA 1079
 QY 432 LeuCYrPheValIleGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluIys 451
 DB 1080 ACTGCAACTGCCACTATCTTATTGATGACAAAAGACACCCACAGATTTTACCAG 1139
 QY 452 SerAspTYrGlyAsnleuThrLeuAlaGluAspThrAsnIleGlySerThrIleuThr 471
 DB 1140 AAGAGATT---CAGCCACAGTAAGAA-----GAGTCAACAGATTAAGTAAC 1190
 QY 472 IleGlnAlaThraAspAlaAspGluProPheThrGlySerSerIySleIleuTYrHisIle 491
 DB 1191 TTAACTGTGTGACGAGATGACCCAGCAACTGAGCATGAGACTGTCTACACTATT 1250
 QY 492 IleIyGlyIyAspSerGlnIyArgIleuGlyValIAspThrAspProIleThraAsnThrGly 511
 DB 1251 ATTAACGAAATCCAGAGGACAGATTGTAATATCATACATCCACAGATTAATGAGGA 1310
 QY 512 TYrValIleIleIySlySlyProIleuAspPheGlnThraIleAlaIyAspSerAsnIleValIle 531
 DB 1311 ATGCTCTGTGTCTCAAACTTTAGACTATGAGATTTCAAGATTTCAACATTTGCGATA 1370
 QY 532 IyAlaIleGluAsnProGluProIleuValIleGlyValIySlyTYrAsnIleAspSerPheAla 551
 DB 1371 AAAGTGAATAATGAAAGCCGTTGATTCAGACATGACCTACCGTCCAGATTCCAGACA 1430
 QY 552 IyPheThrIleuIleValIThrAspValIleAsnGluAlaIleProGlnPheSerGlnIleValIle 571
 DB 1431 ACAGTTCAAGTCAACCGTGAAGATGTAAGAGCCCTGTTTCCACCCAAACCAATG 1490
 QY 572 GlnAlaIyValSerGluAspValAlaIleGlyIThrIyValIyGlyAsnValIThrAlaIyS 591
 DB 1491 ACAGTGAACAAACAAAGAAACATCCATTTGGCAGATGTGTAAACAGTAATGCCACT 1550
 QY 592 AspProGlnIyLeuAsp-----IleSerTYrSerleuArgIyAspThrArgIy 608
 DB 1551 GATCCAGATCTTGTGAACATCAGACTATCAGGTATTTCAAGATTCAGAGACAGC 1610
 QY 609 TrpLeuIyIleAspIleValIThrGlyGlnIlePheSerValIleAspIleAspAspArgIy 628
 DB 1611 TGGCTAGAGATTATCCCAACAAATGATCCGTTGCCACATCTGTCTGTGATGGGA 1670
 QY 629 Ala-----GlySerProTYrArgValIleGlnValAlaIleThrGlnValIleGlyIy 644
 DB 1671 TCTCTCATGTTAGAGATTAACAATACATCTCTCTCTCTGCAATAGACATGTGAAC 1730
 QY 645 SerSerleuSerSerValSerGluPheIleuIleuMetAspValIleAsnAspPro 664
 DB 1731 CTTCTGTCTACAGTACAGAACTTACATCATCCTTGGAGAGCGTCAATGACAAATGTC 1790

QY 665 -----ProArGleuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 1791 CCTCCCTTACCAACACTGGCAAAA-----GTCGTGATGATGCT 1832
 QY 681 SerIaProGlySerLeuLeuIlePheGluAlaThrAspAspGlnHisLeuPheArgGly 700
 Db 1833 AAGATCTCAGAGTAGGTGTAAGTAGAGCATCAGCAAAAGCCCTCCATCCCAACACAGAT 1892
 QY 701 ProHisPheThrPheSerLeuGlySerGlySer-----LeuGlnAsnAspTyrGluValSer 719
 Db 1893 CCA---TTTAATTGTAAGTCAAGTCAAGCAATCTGTCAGAAAGTTATGAGATCAAC 1949
 QY 720 LysIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluArgAla 739
 Db 1950 AAGCTTAACAATACTCAATGCCAGGTTGTCCTGCTTCAA---AACCTGAAAAAGGCCAAT 2006
 QY 740 TyrValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSer 759
 Db 2007 TACAAACATCCCAATCTCAGTACAGATCTTGAAACACCACTTGACTTAACAACACAGAA 2066
 QY 760 LeuProValThrPheCysSerCys 767
 Db 2067 CTGAATTACAGATGCTGTCTCTGC 2090

RESULT 14
 US-08-474-068A-1
 ; Sequence 1, Application US/08474068A
 ; Patent No. 5837525
 ; GENERAL INFORMATION:
 ; APPLICANT: Ranscht, Barbara
 ; TITLE OF INVENTION: T-Cadherin Adhesion Molecule
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Campbell & Flores LLP
 ; STREET: 4370 La Jolla Village Drive, Suite 700
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: United States
 ; ZIP: 92122
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/474,068A
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/213,361
 ; FILING DATE: 14-MAY-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/607,293
 ; FILING DATE: 30-OCT-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Campbell, Cathryn A.
 ; REGISTRATION NUMBER: 31,815
 ; REFERENCE/DOCKET NUMBER: P-LJ 1683
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 535-9001
 ; TELEFAX: (619) 535-8949
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3959 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 45..2181
 ; US-08-474-068A-1

Alignment Scores:
 Pred. No.: 6,22e-59 Length: 3959
 Score: 630.50 Matches: 193
 Percent Similarity: 44.77% Conservative: 115
 Best Local Similarity: 28.05% Mismatches: 283
 Query Match: 14.59% Indels: 97
 DB: 2 Gaps: 25

US-10-025-380-1081 (1-832) x US-08-474-068A-1 (1-3959)

QY 150 AlaThrAspLeuAspAspProAlaThrProAsnGlyGln-----LeuTyrTyrGlnIle 167
 Db 108 GCAGAAAGATTTAAGAA-----TGCACCCCTGGAATTCAGCAAAAGGTTTTTAT 155
 QY 168 ValIleGluLeuProMetIleAsnAsnValMetTyrPheGlnIleAsnAsnValSerThrGly 187
 Db 156 ---ATTGAACAGCCA-----TTTGAATTCACAGAGACCAAGCCA 191
 QY 188 AlaIleSerLeuThrArgIle-----GlySerGlnIleAsn-----ProAlaLys 203
 Db 192 ATTCTGAACCTGGTGTGTTGATGACTGCACAGGGAATTAACAATTGAATTCGAATTCCT 251
 QY 204 AsnProSerTyrAsnLeu-----ValIleSerValIysAspMetGly--- 217
 Db 252 AACCCAGACTTTAAGGTGGAACAGATGATCTTTAGTTGCACTGAAGATGATACGAA 311
 QY 218 -----GlyGlnSerGluAsnSerPheSerAspThrThrSer 229
 Db 312 GCTGGCAGACGCTTTGTTTGTCCATGCACGCTGACAGATGCT---GAGATATATGACGAA 368
 QY 230 ValAspIleIleValThrGluAsnIleTyrPlyAlaProLys----- 243
 Db 369 ATTTGATTGTGGAGCTGATGAGAGACAGACAGATCAATTAAGAAATCTTTAAGATGAA 428
 QY 244 -----ProValGluMetVal 248
 Db 429 GGCACCTTGAATTCACAGCAAAAAAGGCTATTCTGCGCACTCCAAATATTAAATTCGA 488
 QY 249 GluAsnSerThrAspProHisProIleLysIleThrGlnValArgTyrAsnAsp----- 266
 Db 489 GAAATATCAAAAGACCAACCATTTTCCAGATCAAGTTGGCAAGGTCAACGAGTGAAGGACA 548
 QY 267 ProGlyAlaGlnTyrSerLeuValAspLys-----GluLysLeuProArgPheProPhe 284
 Db 549 GAGGAGCAAAAGTCCGACTCTCTGTAAGGAGTATGATCAAGACCCGAAAGAAATTTT 608
 QY 285 SerIleAspGln---GluGlyAspIleTyrValThrGlnProLeuAspArgGluGluLys 303
 Db 609 AGATCAATAGATCAAGTGGGAGTGTCTGTGACCCGACCCCTGGATAGAGAACAAAT 668
 QY 304 AspAlaTyrValPheTyrValAlaValAlaLysAspGluTyrGlyLeuProLeuSerTyrPro 323
 Db 669 GCCAATTATGAGCTGGAAGTTGAAGTAAAGGAAATTAAGTGAATGATGAGCCCA 728
 QY 324 LeuGluIleHisValLysValLysAspIleAsnAspAspProProThrCysProSerPro 343
 Db 729 GTCCGCTTATGATATTTCTGTTATGATCAAAATGATTAAGGCCCATGTTCAAGAGAGA 788
 QY 344 ValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGlyThrLeuThr 363
 Db 789 CCTATGTTGTGCACCTCATGAGAGATCCCTCAAGAAACATGTATGAGGATGACA 848
 QY 364 AlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrTyrGlyIleValGlu 383
 Db 849 GCATTTGATGCTGATGATCTTAGACACAGCAACGCTTTCTCGGTATTAATACCTCAAG 908
 QY 384 GlnThrProLysLeuProMetAspGlyLeuPheLeuGln----- 397
 Db 909 CAGACACCTACCAAACTTCCCAAAATATGTTCTTACATTCAGACCCAGAAAGAGATATT 968
 QY 398 ---ThrTyrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrPro 416
 Db 969 GTACAGAGTGTGTCACTGTACTGCTGATCGTGAACATG-----GAAAGCCG 1019


```

Qy 417 GlnTyrAsnLeuThrIleGluValSerAspLys-----AspPheLysThr 431
Db 1020 AAGTACGAGCTGTTATTAAGCCAGAGATGCGCGCATGATGCGCATTAAGTCTGCA 1079
Qy 432 LeuCyAspPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLys 451
Db 1080 ACTGCACACTCCACTATTCTTATTGATGACAAAACGACCCACGACGATTTTACCAAG 1139
Qy 452 SerAspTyrGlyAsnLeuThrLeuValGluAspThrAsnIleGlySerThrIleLeuThr 471
Db 1140 AAGGAGTGT---CAGGCCACAGATTAAGAA---CGAGTCCAGAGATTAATGATTAAC 1190
Qy 472 IleGlnIleAspIleAspGlnProPheThrGlySerSerLysIleLeuTyrHisIle 491
Db 1191 TTAACTGTTGATGACGAGATGACCCGACCACTGAGCATGAGACGCTGCTTACACTATT 1250
Qy 492 IleGlyAspSerGlnGlyArgLeuGlyValAspThrAspProHisThrAsnThrGly 511
Db 1251 ATTAACGGAATCCAGGCGCAGAGTTTGAATCATCATCCATCCCGACGATTAATGAGGGA 1310
Qy 512 TyrValIleIleLysIleProLeuAspPheGlnThrAlaValSerAsnIleValPhe 531
Db 1311 ATGCTCTGTTGCAAACTTTAGACTATGATTCAGCATTTCCACATTCGATGATA 1370
Qy 532 LysAlaGluAsnProGluProLeuValPheGlyValLysTyrAsnAlaSerSerPheAla 551
Db 1371 AAAGTAAATAATGAAAGCCGTTGATTCAGACATACGCTTCCAGTCCCGATTCACAGCA 1430
Qy 552 LysPheThrLeuIleValThrAspValAsnGlnValProGlnPheSerGlnHisValPhe 571
Db 1431 ACAGTTGAGATCACCGTTGAGAGATGATGAATGAAGCCCTGTTTCCACCCAAACCAATG 1490
Qy 572 GlnAlaLysValSerGluAspValAlaIleGlyThrLysValGlyAsnValThrAlaLys 591
Db 1491 ACAAGTGAACAAACAGAGAAACATCCATATGGCGCATTTGGTTTACAGATTAATGCCACT 1550
Qy 592 AspProGlnGlyLeuAsp-----IleSerTyrSerLeuArgGlyAspThrArgGly 608
Db 1551 GATCAAGATCTTGGCAACATCAGACTATTCAGGATTCAGTTTCAAGATCCAGCAACG 1610
Qy 609 TyrLeuLysIleAspHisValThrGlyLysIlePheSerValAlaProLeuAspArgGlu 628
Db 1611 TGGCTAAGATTAATCCACCAATGATGACGTTGCCACACTGCTGCTCGATGCGGAA 1670
Qy 629 Ala-----GlySerProTyrArgValGlnValValAlaThrGluValGlyGly 644
Db 1671 TCTCTCATGTTTCAGATTAACAAATACACTGCTCTCTCCGCAATAGACAGTGTATAC 1730
Qy 645 SerSerLeuSerSerValSerGluPheHisIleLeuMetAspValAsnAspAsnPro 664
Db 1731 CTTCTCTCATGTTTCAGATTAACAAATACACTGCTCTCTCCGCAATAGACAGTGTATAC 1790
Qy 665 -----ProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
Db 1791 CCTCCCTTAAACCAACACTGCAAAA-----GCTGTGATGATGATGCT 1832
Qy 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspArgGlnHisLeuPheArgGly 700
Db 1833 AAAGATCTCAGAGATGCTGCTACTAGAGCATCAGCAAAAGCTCCCATCCCAACACAGAT 1892
Qy 701 ProHisPheThrPheSerLeuGlySerGlySer-----LeuGlnAsnAspTyrGluValSer 719
Db 1893 CCA---TTTAATATTGAACGATGAGCAATCTGCTCCAGAAAAGTTATGAGATCAAC 1949
Qy 720 LysIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGlnGluArgAla 739
Db 1950 AAGCTTAACATATCTCATGCCAGAGTCTCTGCTCA---AAGCTGAAAAGGCCAAT 2006
Qy 740 TyrValValLeuIleArgIleAsnAspGlyLysArgProProLeuGluGlyIleValSer 759
Db 2007 TACAACATCCCAATCTCAGTACAGATTTCTGAAAACCACTGTGACTAACAACACAGAA 2066

```

```

Qy 760 LeuProValThrPheCysSerCys 767
Db 2067 CTGAATTTACAGCTGTCTCTGCTGC 2090

RESULT 15
US-08-472-481-1
; Sequence 1, Application US/08472481
; Patent No. 5863804
GENERAL INFORMATION:
APPLICANT: Ranscht, Barbara
TITLE OF INVENTION: T-cadherin Adhesion Molecule
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSER: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,481
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/213,361
FILING DATE: 14-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/607,293
FILING DATE: 30-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1686
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-8001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3959 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 45..2181
US-08-472-481-1

Alignment Scores:
Pred. No.: 6,22e-59 Length: 3959
Score: 630.50 Matches: 193
Percent Similarity: 44.77% Conservative: 115
Best Local Similarity: 28.05% Mismatches: 283
Query Match: 14.59% Indels: 97
DB: 2 Gaps: 25

US-10-025-380-1081 (1-832) x US-08-472-481-1 (1-3959)
Qy 150 AlaThrAspLeuAspAspProAlaThrProAsnGlyGln-----LeuTyrTyrGlnIle 167
Db 108 GCGAAGATTTAATA-----TGCACCCCTGCGATTCACGCAAAAGGTTTTTAT----- 155
Qy 168 ValIleGlnLeuProMetIleAsnAsnValMetTyrPheGlnIleAsnAsnLysThrGly 187
Db 156 ---ATTGAACAGCA-----TTGAATTCAACAGACGACGCA 191
Qy 188 AlaIleSerLeuThrArgGlu-----GlySerGlnGluLeuAsn---ProAlaLys 203
Db 192 ATTCTGAACCTGCTGTTGATGATGACGCAAGGGGAATTAACAAATTGAACCTTGAAATTCT 251

```

204 AsnProSerTyrAsnLeu-----ValIleSerValIysAspGlycyl---- 217
 252 AACCCAGACTTTAAGGTGGAACACGATGGATCTTTAGTTGCACCTGAAGAATATATACAGA 311
 218 -----GlyInSerGluAenSerPheSerAspThrThrSer 229
 312 GCTGGCAGAGCTTTTGTTCATCCACGGTCTGACATGCT---GAGGATATGGCAGAA 368
 230 ValAspIleIleValThrGluAsnIleTrypIleProlys----- 243
 369 ATTTCGATTGTTGGACTGATGAGAACGACGATGATTAAGAAATCTTTAAGATAGAA 428
 244 -----ProValGluMetVal 248
 429 GGCAACCTTGGAATTCACAAGACAAAAGGGGTATTCGGGGACTCCAAATATATATTCOA 488
 249 GluAenSerThrAspProIleAspIleThrGlnValArgTrpAsnAsp----- 266
 489 GAAAAATCAAGACCACTTCCTCCGATCACTGTGGCAAGCTCATCAGGAGTGAAGGACA 548
 267 ProGlyAlaGlnTyrSerLeuValAspIys-----GluTysLeuProArgPheProPhe 284
 549 GAGGGAGGAAGTTCGACTCTCTGCTGAAGGAGAGTACAGTCAAGACCCGAAAGAAATTTT 608
 285 SerIleAspGln--GluGlyAspIleTyrValThrGlnProLeuAspArgGluIuLys 303
 609 AGAATCATGATGATCAGTGGGGATGCTCTGTGACCCGACCCCTGGATAGAAACAATTA 668
 304 AspIleTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeuSerTyrPro 323
 669 GCCAATTATGAGCTGGAAAGTTAGATTAACGATTTTAAGTGGGAAATCAATCATGATGGCCCA 728
 324 LeuGluIleHisValIysValIysAspIleAsnAspAsnProProThrCysProSerPro 343
 729 GTCCGCTTACGATATTTCTGTTATTCATCAAAATGATTAACGAGCCGATGTTCAAGAAAGA 788
 344 ValThrValIlePheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGlyThrLeuThr 363
 789 CCTATGTTGGTCACTGATGAGAAAGATCCCTTCACAGAACAACTGATGCGATGACAA 848
 364 AlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArgIleValAlaGlu 383
 849 GCATTGATGCTGATGATCTTCAGCACAGACACGCTCTTCTCGGTTAAACATCTCTCAAG 908
 384 GlnThrProLysLeuProMetAspGlyLeuPheLeuIleGln----- 397
 909 CAGACACCTCAACCAACTTCCCAAAATGTTCTCATATTGACCCGAAAGAGAGATATT 968
 398 ---ThrTyrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrPro 416
 969 GTCAACAGTGGTCACTGTCATCTGTGATGCTGACAGCAATG-----GAAACGGCG 1019
 417 GlnTyrAsnLeuThrIleGluValSerAspLys-----AspPheLysThr 431
 1020 AAGTACAGAGCTGGTATTATGGAAGCCAGAGATATGGCGGCCCATGATGTCGAGCTTACTGGA 1079
 432 LeucysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLys 451
 1080 ACTGCAACTGCCACTATTTCTTATGATACAAAACGACACCCACCGAATTTTATCAACAG 1139
 452 SerAspTyrGlyAsnLeuThrLeuAlaGluAspThrAsnIleGlySerThrIleLeuThr 471
 1140 AAGGAGTT---CAGGCCACAGTAAAGAA-----GAGATGCACAGGATTAATGTAAC 1190
 472 IleGlnAlaThrAspAlaAspGluProPheThrGlySerSerTysIleLeuTyrHisIle 491
 1191 TTAACTGTGGTGAACCGAGATGACCCACCACTGAGAGATGGAAGCTGTCTACATACTATT 1250
 492 IleLysGlyAspSerGluGlyArgLeuGlyValAspThrAspProHisIleThrAsnThrGly 511
 1251 ATTAAACGAAATCCAGGGCAGAGATTGTAATTAACATCAATATCCCAAGACTTAATGAAGGA 1310

OY		512	TyrValIleIleIleLysProLeuAspGluThrAlaIaValSerAsnIleValPhe	531
Db		1311	ATGCCTCTGTGGTCAAACCTTTAGACTATAGATTACGATTTCCACACTTGGCTATA	1370
OY		532	LysAlaGluAenProGluProLeuValPheGlyValIleTryAsnAlaSerSerPheAla	551
Db		1371	AAAGTAGAAAATGAAGAACCCTGTGATTGCCAACATAGCGCTACGGTCCAGTTCCACAGCA	1430
OY		552	LysPheThrIleuIleValThrAspValAsnGluAlaProGlnPheSerGlnHisValPhe	571
Db		1431	ACA GTTCAGATCACCGTTGAGAGATGATTAAGGCCCTGTGTTTTCCACCACCAACCAATG	1490
OY		572	GlnAlaIysValSerGluAspValAlaIleGlyThrLysValGlyAsnValThrAlaLys	591
Db		1491	ACAGTGACAAAAACAAGAACATCCTCATTTGGCAGCACTGGTGTTAACAGTMAATGCCACT	1550
OY		592	AspProGluGlyLeuAsp-----IleSerTySerLeuArgIleAspThrArgGly	608
Db		1551	GATCCAGATACCTTTGGCAACATCAGACTATCAGATTCAGATTTCAGAGATCCAGCAAGC	1610
OY		609	TrpLeuLysIleAspHisValThrGlyGluIlePheSerValAlaProLeuAspArgGlu	628
Db		1611	TGGCTAGGATTATATCCCAACATAGTACCGTTGCCACCACTGGTCTGCTGATGGGAA	1670
OY		629	Ala-----GlySerProTyArgValGlnValAlaIleThrGluValGlyGly	644
Db		1671	TCTCCTCATGTTCCAGATTAACAATATACCTGCTCTTCCGGCAATAGACAGTGGTAA C	1730
OY		645	SerSerLeuSerSerValSerGluPheHisIleuIleLeuMetAspValAsnAspAsnPro	664
Db		1731	CCTCCTGCTACAGGTACAGAACTTTACATCACTTGGAGAGACGTCAATAGACAAATGTC	1790
OY		,665	-----ProArgLeuAlaIleAspTyrThrGlyLeuPhePheCysHisProLeu	680
Db		1791	CCCTCCCTTACCCCAACACTGGCAAAA-----GCTGTGATGATGCT	1832
OY		681	SerAlaProGlySerLeuIlePheGluAlaThrAspAspArgGlnHisIleuPheArgGly	700
Db		1833	AAAGATCTCAGAGTAGTGGTACTAGAGACATCAAGACAAAGACCTTCATCCCAACAGAT	1892
OY		701	ProHisPheThrPheSerLeuGlySerGlySer---LeuGlnAsnAspTrpGluValSer	719
Db		1893	CCA--TTTAAATTGAACTAGTAGAAGCAATCTGGTCCAGAAAAGTTATGGAGATCAAC	1949
OY		720	LysIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluValArgLa	739
Db		1950	AAGCTTAACAATATCATGATGCCCAAGTGTCTGCTTCAA---AACCTGAAAAGGCCAAT	2006
OY		740	TyrValValIleuIleArgIleAsnAspGlyValArgProProLeuGluGlyIleValSer	759
Db		2007	TACAAACATCCCAATCTCAGTACAGATTTCTGGAAAACCACTCTGACTTAACACACAGAA	2066
OY		760	LeuProValThrPheCysSerCys	767
Db		2067	CTGAATTTACAAGTGTCTTCTGC	2090

Search completed: June 22, 2003, 03:24:10
Job time : 217 secs

THIS PAGE BLANK (USPTO)

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2003, 21:24:29 ; Search time 62 Seconds
(without alignments)
1788.137 Million cell updates/sec

Title: US-10-025-380-1081

Perfect score: 4321
Sequence: 1 MIQAMHSLCLMLYATG.....DKGKNVNSAQASEVKEPLRS 832

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A-Geneseq_101002.*

```
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4321	100.0	832	20	AAW74089
2	4321	100.0	832	22	AAW24522
3	908.5	21.0	807	21	AAW99405
4	908.5	21.0	807	22	AAW87574
5	908.5	21.0	807	22	AAW66154
6	908.5	21.0	807	23	AAW83685
7	877.5	20.3	830	22	AAW09959
8	699.5	16.2	906	21	AAW07041
9	699.5	16.2	906	22	ABG14316
10	699.5	16.2	906	23	ABW81474

11	696	16.1	912	23	ABW81471	Chicken N-cadherin
12	694.5	16.1	906	23	ABW57233	Mouse ischaemic co
13	684.5	15.8	916	18	AAW25658	Human cadherin-4.
14	684.5	15.8	916	18	AAW13129	Full length human
15	671.5	15.5	896	15	AAW63533	Human H7-1376 cell
16	671.5	15.5	896	23	AAW78055	Human desmocollin
17	667.5	15.4	847	22	AAW39235	Human polypeptide
18	665.5	15.2	862	22	AAW41021	Human polypeptide
19	657.5	15.2	713	18	AAW25638	Human cadherin-13.
20	657.5	15.2	713	18	AAW13136	Full length human
21	630.5	14.6	712	13	AAW27823	Sequence encoded b
22	630.5	14.6	717	13	AAW27824	Sequence encoded b
23	611	14.1	4555	23	AAW52106	Rat fat 3 protein
24	607	14.0	759	22	ABG21947	Novel human diagno
25	605	14.0	4591	22	ABG22977	Novel human diagno
26	604	14.0	889	23	ABW81472	Mouse E-cadherin p
27	601.5	13.9	3606	22	ABW62595	Drosophila melanog
28	593	13.7	878	15	AAW5060	Sequence of human
29	593	13.7	878	16	AAW85487	Human E-cadherin p
30	593	13.7	878	23	ABW81475	Human E-cadherin p
31	592.5	13.7	117	21	AAW64597	Nonclassical cadhe
32	591.5	13.7	4643	22	ABW71609	Drosophila melanog
33	591	13.7	882	20	AAW09375	Wild-type E-cadher
34	591	13.7	882	21	AAW35730	Human E-cadherin a
35	591	13.7	882	22	AAW73490	Human E-cadherin.
36	591	13.7	882	23	AAW78051	Human E-cadherin p
37	576	13.3	3014	22	AAW68533	Human novel cytol
38	576	13.3	3014	22	AAW02196	Seven-pass transme
39	576	13.3	3028	22	AAW08586	Human NOV7 protein
40	574.5	13.3	796	22	AAW60418	Mouse PrP-binding
41	573.5	13.3	796	15	AAW49730	Sequence encoded b
42	571.5	13.2	788	22	AAW78375	Human protein SEQ
43	571.5	13.2	788	23	ABW57294	Mouse ischaemic co
44	570.5	13.2	840	22	ABG01693	Novel human diagno
45	569.5	13.2	840	22	AAW23903	Human EST encoded

ALIGNMENTS

AAW74089	standard, Protein; 832 AA.
AAW74089;	
04-MAY-1999	(first entry)
Human HPT-1	protein sequence.
Gastro-Intestinal	transport receptor; binding protein; hsi, HPT1;
KW	DBH; hPEPT1; human; GI tract receptor; sucrose-isomaltase complex;
KW	Intestinal peptide-associated transporter; hypertension; diabetes;
KW	osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
KW	therapeutic agent delivery; therapy; probe.
OS	Homo sapiens.
XX	
XX	MO9851325-AA2.
XX	
PD	19-NOV-1998.
XX	
PF	15-MAY-1998; 98WO-US10088.
XX	
PR	15-MAY-1997; 97US-0046595.
XX	
PA	(CYTO-) CYTOGEN CORP.
XX	(ELAN-) ELAN CORP PLC.
XX	
PI	Alvarez VL, Balinka BA, Cagney GM, Carter JM, Lambkin IJ;
XX	Onahony DJ, Patterson CA, Singleton J;
DR	WPI; 1999-009568/01.

DR N-PSDB; AAX18166.

XX New proteins that bind specifically to receptors in the

PT gastro-intestinal tract and related nucleic acid - chimaeras and

PT antibodies, used to deliver therapeutic or diagnostic agents to, or

PT through, the gastro-intestinal tract, e.g. insulin or leuprolide

XX

PS Disclosure; Fig 2; 294pp; English.

XX This sequence is the human HPT-1 protein. The invention relates to

CC purified proteins (I) that bind specifically to at least one of the

CC gastro-intestinal (GI) tract receptors human intestinal

CC peptide-associated transporter (HPT1), hPPT1, D2H and human

CC sucrase-isomaltase complex (hSI). (I) provide active transport of

CC therapeutic agents through human and animal GI tissue (into the blood)

CC for in vivo delivery, particularly for treatment or prevention

CC of hypertension, diabetes, osteoporosis, haemophilia, anaemia, cancer,

CC migraine, or angina pectoris. Specifically they are used to deliver

CC insulin or leuprolide, but many other suitable therapeutic agents are

CC disclosed, including genes or inhibitory nucleic acid, imaging agents and

CC antigens. (I) may also provide targeting to the GI tract. Other uses of

CC (I) are: (i) to determine the level of specified receptors in a sample

CC (in a binding assay); and (ii) to screen for molecules that bind (I).

CC Immunogenic analogues or derivatives of (I) are used to raise antibodies

CC and in immunoassays. The antibodies are used to locate, detect and

CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis

CC etc., also for peptide purification and immobilisation.

XX

XX Sequence 832 AA;

Query Match 100.0%; Score 4321; DB 20; Length 832;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 832; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MILQAHLSLCLMLYLATGKGEGKFGPKPMFTFSYEGQSPQIIFQKAPPAVTF 60

DB 1 MILQAHLSLCLMLYLATGKGEGKFGPKPMFTFSYEGQSPQIIFQKAPPAVTF 60

QY 61 ELTGETDNI FVIEREGGLLYNRALDRETRSTHNIQVAALDANGIIIEGVPDITTEVKDIN 120

DB 61 ELTGETDNI FVIEREGGLLYNRALDRETRSTHNIQVAALDANGIIIEGVPDITTEVKDIN 120

QY 61 ELTGETDNI FVIEREGGLLYNRALDRETRSTHNIQVAALDANGIIIEGVPDITTEVKDIN 120

DB 61 ELTGETDNI FVIEREGGLLYNRALDRETRSTHNIQVAALDANGIIIEGVPDITTEVKDIN 120

QY 121 DNRPTFLOSKEYGSRQNSRPGKPELYNATDLDPAFNCQLYQVITQIPMINNWF 180

DB 121 DNRPTFLOSKEYGSRQNSRPGKPELYNATDLDPAFNCQLYQVITQIPMINNWF 180

QY 121 DNRPTFLOSKEYGSRQNSRPGKPELYNATDLDPAFNCQLYQVITQIPMINNWF 180

DB 121 DNRPTFLOSKEYGSRQNSRPGKPELYNATDLDPAFNCQLYQVITQIPMINNWF 180

QY 181 QINNKTGAI SLTREGSOELNPAKNSYNLVISVKMGQSGNSPSDTTSVDIIVTENIMK 240

DB 181 QINNKTGAI SLTREGSOELNPAKNSYNLVISVKMGQSGNSPSDTTSVDIIVTENIMK 240

QY 181 QINNKTGAI SLTREGSOELNPAKNSYNLVISVKMGQSGNSPSDTTSVDIIVTENIMK 240

DB 181 QINNKTGAI SLTREGSOELNPAKNSYNLVISVKMGQSGNSPSDTTSVDIIVTENIMK 240

QY 241 APKEVENVENSTDPHPKITQVRANDPGAQYSLVDKELPPRPSIDEGSITVYTOPIDR 300

DB 241 APKEVENVENSTDPHPKITQVRANDPGAQYSLVDKELPPRPSIDEGSITVYTOPIDR 300

QY 241 APKEVENVENSTDPHPKITQVRANDPGAQYSLVDKELPPRPSIDEGSITVYTOPIDR 300

DB 241 APKEVENVENSTDPHPKITQVRANDPGAQYSLVDKELPPRPSIDEGSITVYTOPIDR 300

QY 301 EEKQAYVYAAKDEYGPPLSYPLEIHVKVDINDNPTCSPVTFVQENERLGNISIG 360

DB 301 EEKQAYVYAAKDEYGPPLSYPLEIHVKVDINDNPTCSPVTFVQENERLGNISIG 360

QY 301 EEKQAYVYAAKDEYGPPLSYPLEIHVKVDINDNPTCSPVTFVQENERLGNISIG 360

DB 301 EEKQAYVYAAKDEYGPPLSYPLEIHVKVDINDNPTCSPVTFVQENERLGNISIG 360

QY 361 TLTAHADRDEENTANSFLNRYIVEQTPKLPMDGLFIQTYAGMLQIAKOSLKKDTPQYNL 420

DB 361 TLTAHADRDEENTANSFLNRYIVEQTPKLPMDGLFIQTYAGMLQIAKOSLKKDTPQYNL 420

QY 361 TLTAHADRDEENTANSFLNRYIVEQTPKLPMDGLFIQTYAGMLQIAKOSLKKDTPQYNL 420

DB 361 TLTAHADRDEENTANSFLNRYIVEQTPKLPMDGLFIQTYAGMLQIAKOSLKKDTPQYNL 420

QY 421 TIEVSDKQFTLCPQIVINDINOIPIFEKSDVGNLTLADTNIGTILITQATDDEP 480

DB 421 TIEVSDKQFTLCPQIVINDINOIPIFEKSDVGNLTLADTNIGTILITQATDDEP 480

QY 421 TIEVSDKQFTLCPQIVINDINOIPIFEKSDVGNLTLADTNIGTILITQATDDEP 480

DB 421 TIEVSDKQFTLCPQIVINDINOIPIFEKSDVGNLTLADTNIGTILITQATDDEP 480

QY 481 FTGSSKILYHITIKDSEGRGLGVDTPHNTGTVIIRKPLDEETAANSIVKRAENPEPLV 540

DB 481 FTGSSKILYHITIKDSEGRGLGVDTPHNTGTVIIRKPLDEETAANSIVKRAENPEPLV 540

QY 481 FTGSSKILYHITIKDSEGRGLGVDTPHNTGTVIIRKPLDEETAANSIVKRAENPEPLV 540

DB 481 FTGSSKILYHITIKDSEGRGLGVDTPHNTGTVIIRKPLDEETAANSIVKRAENPEPLV 540

QY 541 FGKYNASSPAKFTLITVDVNEAPOFSQHVFOAKVSEDAVIGTKVGNATADPEGLDISY 600

DB 541 FGKYNASSPAKFTLITVDVNEAPOFSQHVFOAKVSEDAVIGTKVGNATADPEGLDISY 600

QY 601 SLRGTRGMLKIDHVTGEIFSVAPLDREAGSPYRVQVAVATEVGGSSLSVSEPHLLMDV 660

DB 601 SLRGTRGMLKIDHVTGEIFSVAPLDREAGSPYRVQVAVATEVGGSSLSVSEPHLLMDV 660

QY 661 NDNPRLAKYTGTLFPHPLSAGSLIFEAITDDQHLFRGPHFTFSLGSSSLONDEWVSK 720

DB 661 NDNPRLAKYTGTLFPHPLSAGSLIFEAITDDQHLFRGPHFTFSLGSSSLONDEWVSK 720

QY 721 INGTARLSRPHDPEERAVVVLIRINDGSRPPLGIVSLPVPFCSCVEGSCFRPAGHOT 780

DB 721 INGTARLSRPHDPEERAVVVLIRINDGSRPPLGIVSLPVPFCSCVEGSCFRPAGHOT 780

QY 781 GIFTGVAGVIGLLTTLVLIGIILAVVFIIRIKDKGKNVESAQASEVKPIRS 832

DB 781 GIFTGVAGVIGLLTTLVLIGIILAVVFIIRIKDKGKNVESAQASEVKPIRS 832

RESULT 2

AAM24522

ID AAM24522 standard; Protein; 832 AA.

XX

AC AAM24522;

DT 12-OCT-2001 (first entry)

XX

DE C980P similar amino acid sequence (GENESEQ W740898).

XX

KM Human; immunotherapy; diagnosis; colon cancer; colon tumour;

KW immunogenic; gene therapy; vaccine; colonic cancer.

XX

OS Homo sapiens.

XX

EN WO200149716-A2.

PD 12-JUL-2001.

XX

PF 29-DEC-2000; 2000WO-US55596.

XX

PR 30-DEC-1999; 99US-0476296.

PR 10-JAN-2000; 2000US-0480321.

PR 15-FEB-2000; 2000US-0504629.

PR 06-MAR-2000; 2000US-0519444.

PR 19-MAY-2000; 2000US-0575251.

PR 29-JUN-2000; 2000US-0609448.

PR 28-AUG-2000; 2000US-0649811.

XX

PA (CORI-) CORIXA CORP.

XX

PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;

PI King GE, Wang T, Jiang Y;

XX

DR WPI; 2001-441847/47.

XX

PT Colon tumor associated proteins and nucleic acids useful for the

PT prevention, diagnosis and treatment of colonic cancer -

XX

PS Claim 2; Page 469-472; 472pp; English.

XX

CC The present invention describes colon tumour associated proteins (I) and

CC the polynucleotides (II) that encode them. (I) have cytostatic activity.

CC (I) and (II) can be used in gene therapy and vaccine production. (I) and

CC (II) may be used in the prevention, diagnosis and treatment of diseases

CC associated with inappropriate colon tumour associated protein (TCAP)

CC expression, such as colonic cancer. For example, (I) and (II) may be

CC used to treat disorders associated with decreased expression by

CC rectifying mutations or deletions in a patient's genome that affect the

CC activity of TCAPs by expressing inactive proteins or to supplement the

CC patients own production of them. Additionally, (II) may be used to

CC produce the TCAP proteins, by inserting the nucleic acids into a host

CC cell culturing the cell to express the protein. (II) and its

CC complementary sequences may also be used as DNA probes in diagnostic

CC polymerase chain reaction (PCR) and hybridisation assays to detect and


```

Db      684 LGNPTVQDRMRQLTNGSHAYLTTLALHWEPREHIIIPVAVSHNAQ---MMQLVRYIVC 740
Qy      766 SC-VEGSCFRPAGHOTGIPVGNVAVGILLTTLVIGIILAVFIRIKDKGKNVESAOA 824
Db      741 RCNVEGQCRKVRGMKMPITLSAVGILVGTVAIGIFILIFTHWTMSRKDDPDQPADS 800
Qy      825 SEVK 828
Db      801 VPLK 804

RESULT 4
AAB87574
ID      AAB87574 standard; Protein; 807 AA.
AC      AAB87574;
XX
DT      15-MAY-2001 (first entry)
DE      Human PRO1340.
XX
KM      Human; PRO protein; mapping.
OS      Homo sapiens.
XX
PN      WO200116318-A2.
PD      08-MAR-2001.
PF      24-AUG-2000; 2000WO-US233328.
XX
PR      01-SEP-1999; 99WO-US20111.
PR      15-SEP-1999; 99WO-US21090.
PR      07-DEC-1999; 99US-0169485.
PR      09-DEC-1999; 99US-0170262.
PR      11-JAN-2000; 2000US-0175481.
PR      18-FEB-2000; 2000WO-US04341.
PR      22-FEB-2000; 2000WO-US04342.
PR      01-MAR-2000; 2000WO-US05601.
PR      03-MAR-2000; 2000US-0187202.
PR      25-APR-2000; 2000US-0193937.
PR      22-MAY-2000; 2000WO-US14042.
PR      05-JUN-2000; 2000US-0209832.
XX
PA      (GETH ) GENENTECH INC.
PI      Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ,
PI      Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;
XX
DR      WPI; 2001-183260/18.
DR      N-PSDB; AAF92106.
XX
PT      Eighty four nucleic acids encoding PRO polypeptides, useful in
PT      molecular biology, including use as hybridization probes, and in
PT      chromosome and gene mapping.
XX
PS      Claim 12; Fig 98; 278pp; English.
XX
CC      The present sequence is a human PRO polypeptide (secreted and
CC      transmembrane). The PRO protein, and PRO agonists, PRO antagonists or
CC      anti-PRO antibodies are useful for preparation of a medicament useful in
CC      the treatment of a condition which is responsive to the PRO protein,
CC      agonists, antagonists or anti-PRO antibodies. The PRO protein may also be
CC      employed as molecular weight markers for protein electrophoresis. The PRO
CC      coding sequence has applications in molecular biology, including use as
CC      hybridisation probes, and in chromosome and gene mapping.
XX
SO      Sequence 807 AA;
XX
Query Match 21.0%; Score 908.5; DB 22; Length 807;
Best Local Similarity 30.2%; Pred. No. 1.le-64;

```

```

Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;
Qy      62 LTGE-----FDNIVVERE-GLLYNRALDETSSTNLOVAALDANGIIVEGVPTREV 116
Db      57 LSGDSGKATRGPRAMPDPSGFLLYTRALDREDAEYQLOVLTLEMOSGHVIMGQPVLVHV 116
Qy      117 KDINDNRPTFLQSGRYGVSARONSRPGKPELYVNATLDDPATNGOLYYOIVLOPMINN 176
Db      117 KDENDQVPHFSQALYRARLSRGTRPGIFLELASRDEGCTANSDLRPHILLSQAAPQPS 176
Qy      177 VMTFOINNTGALSLTRGSGQELNPAKNSYNLVISKMGSGSENSFSDTSVDIIVTE 236
Db      177 PDWFOLEPRIGALALPKGSTSLDHALERTYQLLVQVKMGDQA-SGHQATATVAVSIIIE 235
Qy      237 NIKAKPKPVEMVNSTNDPHPIKITQYRMNDPGAQYSIVDEKULPRPFPSIDOGSDIYVTO 236
Db      236 STWVSLPEPIHLAENLKVLPHMAQVHWGSGDVHYHL---ESHPPGPFVNAAGNLVYTR 232
Qy      297 PLDREKDAYVFAVAKDEYKPLSYPLIHWVKDINDNPPTCPSPVTVFEVQENERLG 356
Db      293 ELDRBAQAEYLLQVRAQNSHGEDIYAAPLELHVLMENDNVPICPRPDPVSIPELSPPG 352
Qy      357 NSIGITLHADRDSEKNTANSFLNRYIVEQTPKLEMDG-LFLIQTYAGMLQAKOSLKK-QD 414
Db      353 TEVTRLSAEDADAPGSPNSHVYQLSPSEDEVEGRACVDPSTSGSVTLGVPLRAGQN 412
Qy      415 TPQYNLTIEVSDKD--FKTLCFQVINVIDINDOIPIFEKSDVGNLTABDTNGISLTIT 472
Db      413 ILLVLTAMDLAGAGGFSSTCEVEVAVTDDNDAPFITSQIGPISLPEDVEBGTIVAML 472
Qy      473 QATDAD-EPFTGSSKIL-YHIIKDGSEGRIGVDDPHTNTGVYIIR-KPLDETAVASN 528
Db      473 TAITDADLEP---AFRLMDFAIERGDTEGFGDMEP--DSGHVRLALCKNLSYEAPSHSE 527
Qy      529 IYFKAENPEPLVFGVKYNNASPFKFTLIYTDVNEAQFSQHVQAVSEDAIGTKYGVN 588
Db      528 VVVVQSVAKLV-GPSPGCGATATVTLVERWVPPTKLOESYEASVPI SAPAGSEFLIT 586
Qy      589 TAKDPGGLDISYLRDTRGWLKIDHVTGEIPEVAFLD-REAGSPRYOVVATEVSGSSL 647
Db      587 QPSDPISRITRPSLVNDSGMLCEKFSGVHQAQSLQAQPDYITVLEADDTA---- 642
Qy      648 SSVSEFHLILMDVNDNPPRIAKDYTLFFCHPLSAPGSLIF-EATDDOHLFFGPHFTFS 706
Db      643 -----LTLAPVPSQ-----YLCTRPDHGLIVSGSKPDLASGHP-YSTFT 683
Qy      707 LG-SGSLQNDMEYSKINGTHARLSTRHTDPEERAVVYLIRINDGRRPLEGIYSLPVTF 765
Db      684 LGNPTVQDRMRQLTNGSHAYLTTLALHWEPREHIIIPVAVSHNAQ---MMQLVRYIVC 740
Qy      766 SC-VEGSCFRPAGHOTGIPVGNVAVGILLTTLVIGIILAVFIRIKDKGKNVESAOA 824
Db      741 RCNVEGQCRKVRGMKMPITLSAVGILVGTVAIGIFILIFTHWTMSRKDDPDQPADS 800
Qy      825 SEVK 828
Db      801 VPLK 804

RESULT 5
AAB66154
ID      AAB66154 standard; protein; 807 AA.
AC      AAB66154;
XX
DT      02-APR-2001 (first entry)
DE      Protein of the invention #66.
XX
KM      Secreted; transmembrane; gene therapy.
OS      Unidentified.
XX

```


QY 444 DOIPIF-EKSDYGNLTLAEDNTIGSTLTITQATDADPEFTGSSKILYHIKDGSEGLGV 502
 DB 377 DNPEFTAMTFYGEV---PENRVDIIYANLTVTDKQDPHPANNAYVRISGDPFGFAI 433
 QY 503 DTDPHNTGYVYIIKKPLDFETAASNIVFKAENPEPLVFGVKNASSPAKFTLLVTDVNE 562
 DB 434 QTDPNNDGLVTVVKPIDEFETNMFVLTVAAENQVPLAKGIQHPPOSTATVSVTVIDVNE 493
 QY 563 APOGSHVFOAKSSEDAIGTKVGNVTAQDBEG---DISYSLRGDTRGMKIDHYTGEI 619
 DB 494 NPYFAPNPKIIRQEEGHAGTMTLTTFAODPDRMOONIRYTKLSDPANMLKIDPVANGOI 553
 QY 620 FSVAPLDREA---GSPYRQOVVATEVGGSSLSVSEFHLLMDVNDNPPRLAKDYTGFLF 675
 DB 554 TTITAVLDRESPPANNKNNIYATFLASDNGIPRMSGTGTLQIYLLDINDMAQV----- 606
 QY 676 FCHPLSA-----PGSLIFEATDDQHLFRGRH-FTFSIGSSGLQNDWEVSKINGTHAR 727
 DB 607 ---POEATCETPDPSNINITALDYDIDPNAGPFAFDLPSPVITIKRWITTRLNGDPAQ 663
 QY 728 LSTRHDPFERAVVLRINDGGRPPLEGVLSLPVTCSC-VGSSCFRPAHQGTGFTV- 785
 DB 664 LNLKIFLEAGIYEVPIITDTSNMPKSNISILRVKVCQCCDSNGDC-----TDVDRIV 716
 QY 786 --GMAVGILLTTLVIGIILAVFIRIKKDKGKNVESAQAQSEVKP 829
 DB 717 GAGLGTAITAILCTIITLITLILVFMVMMKRKRKROAKOLIDP 762

RESULT 10
 ABB81474
 ID ABB81474 standard; Protein; 906 AA.
 AC ABB81474;
 XX 30-AUG-2002 (first entry)
 DT
 XX
 DE Human N-cadherin protein SEQ ID NO:46.
 KM Cadherin; alpha-catenin; cancer; beta-catenin binding domain; melanoma;
 XX o-catenin; colon cancer.
 OS Homo sapiens.
 XX
 PN US2002045591-A1.
 PD 18-APR-2002.
 PF 17-JUL-2001; 2001US-0905963.
 PR 26-MAY-1998; 98IL-0124650.
 PR 26-MAY-1999; 99US-0318633.
 XX
 PA (GEIG/) GEIGER B.
 PA (BENZ/) BEN-ZE'EV A.
 PA (SADO/) SADO E.
 XX
 PI Geiger B, Ben-Ze'ev A, Sado E;
 XX
 DR WPI; 2002-499105/53.
 DR N-PSDB; ABN89391.
 XX
 PT New construct encoding soluble cytoplasmic portion of cadherin
 PT including beta catenin binding domain useful in treating cancer
 PT associated with high beta-catenin activity e.g. colon cancer and
 PT melanoma
 XX
 PS Example 3; Page 45-48; 102pp; English.
 CC The present invention describes a pharmaceutical composition for treating
 CC cancer associated with abnormally high beta-catenin activity. The
 CC pharmaceutical composition comprises a gene therapy vehicle harbouring a

CC polynucleotide that contains: (1) a nucleotide sequence encoding a
 CC soluble cytoplasmic portion of a cadherin which lacks a transmembrane
 CC portion and an extracellular portion of the cadherin, and includes a
 CC beta-catenin binding domain; and (b) an upstream promoter for directing
 CC expression of the soluble cytoplasmic portion of the cadherin in a
 CC mammalian cell. Also described is a pharmaceutical composition for
 CC treating cancer associated with abnormally high activity levels of
 CC beta-catenin comprising a gene therapy vehicle harbouring a
 CC polynucleotide that contains: (a) a nucleotide sequence encoding an
 CC o-catenin; and (b) an upstream promoter for directing expression of the
 CC o-catenin in a mammalian cell. The pharmaceutical compositions have
 CC cytostatic activity and can be used in the suppression of
 CC beta-catenin-mediated transactivation. They can be used for treating
 CC cancers associated with abnormally high activity levels of beta-catenin
 CC such as colon cancers and melanomas, by reducing these high activity
 CC levels of beta-catenin in mammalian cells. The present sequence
 CC represents human N-cadherin which is used in the exemplification of the
 CC present invention.

CC Sequence 906 AA;
 XX

Query Match 16.2%; Score 699.5; DB 23; Length 906;
 Best Local Similarity 27.4%; Pred. No. 1.3e-47;
 Matches 210; Conservative 122; Mismatches 331; Indels 103; Gaps 21;

QY 117 KQINDNRPFLQSGYGVKQNSRPGKFLYVNAVTDLPATPQGLYYQIVIQLPIMNN 176
 DB 47 KDVBEGQ-LLANKF-----SNCGKRRKQYESSSEPADFVDDGVY-AVNSFPLSSE 98
 QY 177 VMYFOINKKTGAISTRESQELINPAKNSYLVISVKMGSGGSENSFSDTTSVDIIV- 234
 DB 99 HAKFLIY-----AQDKETQEKQVAVKLSLKP--LTSESVSASAEVEELVFP 144
 QY 235 -----TENIWKAPKPYEMTENSIDPHIKITQVRN-----DPGAQ 270
 DB 145 RQFSKSHGLQROKRWVIP-PTNLPENSGRFPQELVTRSDRDKNLSIRYVTGEGA- 202
 QY 271 YSLVDEKELRPFPFSIDQ-EGDIYVTOPLDREKDAVFAVAKDEYKSLSPLEIHVK 329
 DB 203 -----DQPTGTIINIPISQGLSVTKPLDREGIARFLRAHVDINGQVENPIDIVIN 256
 QY 330 VKDINDNPTCPSPVTFVEQENERLGNISIGTLTAHDBEENTANSFLNRYIEQTPKLP 389
 DB 257 VIDMNDNRPFLHQVWNGVTFEGSKPGTYVMVTTAIDADPNALMGLRARIYVSOAHPSTP 316
 QY 390 MDGLFLIQYAGMLQAKQSLKKQDTPQYNLTIEVSDK-----FKTLCFVQINVIDIN 443
 DB 317 SPNNFTINNETGDIITVAAGLDREKVOQYTLIIQATDMEGIPTYGLSNATATAVITVDN 376
 QY 444 DOIPIF-EKSDYGNLTLAEDNTIGSTLTITQATDADPEFTGSSKILYHIKDGSEGLGV 502
 DB 377 DNPEFTAMTFYGEV---PENRVDIIYANLTVTDKQDPHPANNAYVRISGDPFGFAI 433
 QY 503 DTDPHNTGYVYIIKKPLDFETAASNIVFKAENPEPLVFGVKNASSPAKFTLLVTDVNE 562
 DB 434 QTDPNNDGLVTVVKPIDEFETNMFVLTVAAENQVPLAKGIQHPPOSTATVSVTVIDVNE 493
 QY 563 APOGSHVFOAKSSEDAIGTKVGNVTAQDBEG---DISYSLRGDTRGMKIDHYTGEI 619
 DB 494 NPYFAPNPKIIRQEEGHAGTMTLTTFAODPDRMOONIRYTKLSDPANMLKIDPVANGOI 553
 QY 620 FSVAPLDREA---GSPYRQOVVATEVGGSSLSVSEFHLLMDVNDNPPRLAKDYTGFLF 675
 DB 554 TTITAVLDRESPPANNKNNIYATFLASDNGIPRMSGTGTLQIYLLDINDMAQV----- 606
 QY 676 FCHPLSA-----PGSLIFEATDDQHLFRGRH-FTFSIGSSGLQNDWEVSKINGTHAR 727
 DB 607 ---POEATCETPDPSNINITALDYDIDPNAGPFAFDLPSPVITIKRWITTRLNGDPAQ 663
 QY 728 LSTRHDPFERAVVLRINDGGRPPLEGVLSLPVTCSC-VGSSCFRPAHQGTGFTV- 785
 DB 664 LNLKIFLEAGIYEVPIITDTSNMPKSNISILRVKVCQCCDSNGDC-----TDVDRIV 716

QY 786 --GMAVGILTLTLVIGIILAVFIRIKDKGDNVESQAQSEVKP 829
 DB 717 GAGLGTGAIIAIIILCIITLILVLMFVMMKRRDKERQAKQLLIDP 762

RESULT 11

ABB81471
 ID ABB81471 standard; Protein; 912 AA.

AC ABB81471;

DT 30-AUG-2002 (first entry)

DE Chicken N-cadherin protein SEQ ID NO:2.

KM Cadherin; alpha-catenin; cancer; beta-catenin binding domain; melanoma;
 KM o-catenin; colon cancer.

OS Gallus gallus.

PN US2002045591-A1.

PD 18-APR-2002.

PF 17-JUL-2001; 2001US-0905983.

PR 26-MAY-1998; 98IL-0124650.

PR 26-MAY-1999; 99US-0318633.

PA (GEIG/) GEIGER B.

PA (BENZ/) BEN-ZE'EV A.

PA (SADO/) SADO E.

PI Geiger B, Ben-Ze'ev A, Sado E;

XX MPI, 2002-499105/53.

XX N-PSDB; ABB89350, ABB89351.

PT New construct encoding soluble cytoplasmic portion of cadherin

PT including beta catenin binding domain useful in treating cancer

PT associated with high beta-catenin activity e.g. colon cancer and

PT melanoma

PS Example 1; Page 19-21; 102pp; English.

XX The present invention describes a pharmaceutical composition for treating
 CC cancer associated with abnormally high beta-catenin activity. The
 CC pharmaceutical composition comprises a gene therapy vehicle harbouring a
 CC polynucleotide that contains: (1) a nucleotide sequence encoding a
 CC soluble cytoplasmic portion of a cadherin which lacks a transmembrane
 CC portion and an extracellular portion of the cadherin, and includes a
 CC beta-catenin binding domain; and (b) an upstream promoter for directing
 CC expression of the soluble cytoplasmic portion of the cadherin in a
 CC mammalian cell. Also described is a pharmaceutical composition for
 CC treating cancer associated with abnormally high activity levels of
 CC beta-catenin comprising a gene therapy vehicle harbouring a
 CC polynucleotide that contains: (a) a nucleotide sequence encoding an
 CC o-catenin; and (b) an upstream promoter for directing expression of the
 CC o-catenin in a mammalian cell. The pharmaceutical compositions have
 CC cytostatic activity and can be used in the suppression of
 CC beta-catenin-mediated transactivation. They can be used for treating
 CC cancers associated with abnormally high activity levels of beta-catenin
 CC such as colon cancers and melanomas, by reducing these high activity
 CC levels of beta-catenin in mammalian cells. The present sequence
 CC represents chicken N-cadherin which is used in the exemplification of
 CC the present invention.

XX Sequence 912 AA;

Query Match 16.1%; Score 696; DB 23; Length 912;
 Best Local Similarity 26.9%; Pred. No. 2.6e-47;
 Matches 209; Conservative 117; Mismatches 328; Indels 124; Gaps 22;

QY 126 FLQSKTEGVSVRNSRREKPELYVNAATLDDPATPNQOLYYQIVIOLEMINNMYFOIN-- 183
 DB 41 FPDVSAVVSRSVHGQPLNVRFGSCDE-----NRKIFYGSGSEP 81

QY 184 ----NKTGAI SLTREGSOELNPAKPSYLVISVDMGQS-----ENSRSDTT 228

DB 82 EDRFVGEDGVVAERFQLSAEPTL-----FVVSARDKETQEBWMQKVLTPPAFAGS 136

QY 229 SVDIIVTENT-----WKAKPPEVMVENSSTDPHPKIKITQVRMNDPGA 269

DB 137 EKQOKKIEDIIFWQYQKSSHLKQKRDWVTP-PTNLNPNSSGPPFOELVRLR----- 189

QY 270 QVSLVDKEXLPPPPFSI-----DQ-----EGDIYVTOPLDREKDAVYFAVAK 313

DB 190 ----SDRDKSLRSYVTGPGADQPPGIFITINPISGQSLVTPKLDREQLASFLRHAAV 245

QY 314 DEYCKPLSYPLLEHVKKYKQINDNPPCPSPVYVFEVQENRCLNSIGTLTAHDDDEBNTA 373

DB 246 DVNGQVNEPIDIVINVIDMNDNRPEFLHQVWNGTPEGSKPGTYVMTVAIDADDPNAQ 305

QY 374 NSFPLNRYVEQTEKLPMDGLFLQTYAGMLQAKOSLKKQDTPQYNLTLEVSXD----- 428

DB 306 NGMLRIRILSQABSSSPNNFTINNETGDIITVAAGLDREKVOQYTLITQATMEGNPTY 365

QY 429 -FKTLGFOVINVIDINDQIPIF-EKSDYGNLTLAEDTNIGSTILITQATDADPEPTGSSK 486

DB 366 GLSNTADAVITVVDVNDNPEFTAMTFYGSV---PENRDVIVANLTVTDKDDPHIPAMN 422

QY 487 ILYHIITKDSSEGLGVDTPDHTNTGYVILIKPLDFTFAVSNIVFAQENPEPLVFGVKIN 546

DB 423 ARYQMTGSDPTGQFTILTPDNSNDGLVTVKRPIDFETNRMFVLTVAENQVPLAKGIQHP 482

QY 547 ASSEFAKFTLVTDVNEAPQSOHVFOAKVSEDAVIGTKGNVYAKXP-----EGLDISYSL 602

DB 483 POSTATVSVITVIDNNSPTFVPPKLVROBEGGLASSMLTTFARDPRYMQGTSLSYSK 542

QY 603 RGDTRGMLKIDHYTGIFSVAPLDREA---GSPYRVQVATEVGGSSLSVSEPHILIM 658

DB 543 LSPDPAWMLKIDPNQGITTTAVLDRESIYQNNMVAATPLASONGIPPMAGGTGTLQIYLL 602

QY 659 DVNDNPPRL-AKDYTLFFCHPLSAPGSLIFEATDDOHLFRGPHFTFSISGG--SLOND 715

DB 603 DINDNAPQVNPKEAT--TCETLQ-DNAINITAVDDPIDPNAGP-PAFELPDSPPSISKRN 657

QY 716 MEVSKINGHARLSTHTDPEERAYVVLIRINDGSRPPLEGIVSLVTPFSC--VEGSCFR 774

DB 658 WTVIRISGHAQLSLRPLRLEAGIYDVPVITDSGNPHASSTSVLKVYCCQCDINDGC-- 715

QY 775 PAGHQGTPIPV--GMAVGILTLTLVIGIILAVFIRIKDKGDNVESQAQSEVKP 829

DB 716 ----TDVDRIVAGAGTGTAIIAIIILCIITLILVLMFVMMKRRDKERQAKQLLIDP 768

RESULT 12

ABB57233
 ID ABB57233 standard; Protein; 906 AA.

AC ABB57233;

DT 07-MAR-2002 (first entry)

DE Mouse ischaemic condition related protein sequence SEQ ID NO:606.

KM Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 KM vasospastic ischaemia; ischaemic condition; ischaemic disease.

OS Mus musculus.

PN MO200188188-A2.

PD 22-NOV-2001.

PF 18-MAY-2001; 2001MO-JP04192.

```

XX 18-MAY-2000; 2000JP-0145977.
PR (UNY1-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX Iahikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX WPI; 2002-034733/04.
DR N-PSDB; AB199593.
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
XX Claim 2; Page 1560-1564; 2690PP; English.
PS
XX The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (AB199202 to AB199912, encoding
CC the protein sequences in AB57020 to AB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischaemic condition-improving
CC drugs or therapeutics for ischaemic diseases. AB199913 and AB199914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.
XX
SQ Sequence 906 AA;
Query Match 16.1%; Score 694.5; DB 23; Length 906;
Best Local Similarity 27.6%; Pred. No. 3.4e-47;
Matches 211; Conservative 122; Mismatches 332; Indels 99; Gaps 23;
QY 117 KDINDNPTPLQSGKESVQNSRQKPELYVNAITDDPATRPGOLYQVIOLOPMINN 176
DB 47 KDVEGGP-LLNVFNSCNKR-----RKVOYSESPADFKVDEGTVY--AVRSFPLTAE 98
QY 177 VMYFOI---NNKIQ-----AISTREGSOELINPAKNP-SYVLVTSVKDM--GSGSENSF 224
DB 99 QAKLTIYADKETEOKQAVNLSREPTLTPPEKHEHIEIFYPPQOLAKHSGALROK 158
QY 225 SDTTSVDIIVTENIWKAPKPEVWENSTDPHPKITOVNAN-----DPGAQYS 272
DB 159 RD-----WVIP-PINLPENSRGPPQELVIRSDRDKNLSLYSVTGPQA--- 202
QY 273 LVNKEKLPFPFESIDQ-EGDIYVQPLDREKDAVVFVAVKDYGKPLSPLFIHKVK 331
DB 203 ----DQPTGIFILNPIISGQLSVTKPLDRELIAFPHLRAHAVDINGQVENPIDIVINVI 258.
QY 332 DINDNPTCSPVTVFEVOENRNLGNSIGTLTAHDREENTANGFLNRYIEQTPKIPMD 391
DB 259 DANDNRBEFLHQVWNGSVPESSKPGTYMTVTAIDADDPAALNMLKTRILISQAPSPSP 318
QY 392 GLPLIQTAYGMLQAKOSLKKQDTPQYNLTIEVSKD-----FKTLCVOQINVIDINDQ 445
DB 319 NMFTINNETGDIITVAGLDRKVOQVTLTIQATDMEGNPTYGSLNTATAVITVTVNDN 378
QY 446 IPIF-EKSDVGNLTIAEDTNGSTITLTQATDADBEPTGSSKILYHIKGSBERLQVDT 504
DB 379 PPEFTAMTFYGEV---DENRVDIVANMLTVTDKQPHIPANNAAYRISGGPPTRFALTT 435
QY 505 DPHNTGTVTIKKPLDEFETAIVSNIVFEKAPNEPELVFGVKNASFPKFTLIIVDVNEAP 564
DB 436 DPNNDGLIVYVKKRIDETNRMFVLTAAENQVFLAKGIQPPQSTATVSTVIVDVENP 495
QY 565 QFSCHVFOAKVSEDAVIGTKVGNVTAQPEGL---DISYSLRGDTRGMLKIDHTVGEIFS 621
DB 496 YFAFNPKIRQEGHAGTMTLTTLTAQDPDRYMOONIRYTLSDPAWMLKIDPVGQITTT 555

```

```

QY 622 VAPLDREA-----GSPRYGVAVTEVGSSLSVSEFHLIMDVNDNPPRLAKDYTLGFC 677
DB 556 IAVLDRESRPVQNNITNATFLASDNGIPKMSGCTGTYIHLDDINDNAPQVL----- 606
QY 678 HPLSA-----PGSLIFPATDDQHLFRGPH-FTFSLGSGSLQNDWEVSKINGTARLS 729
DB 607 -PQEAETCETPEBNSINIAALDYDIDPNAGPFPAFDLPISPVTKRWMTINRLNGDPAQLN 665
QY 730 TRHTDPEERAYVYLITINQGRPRELEGYSLPYTPFSC-VEGSCFRPAGHGTGIPVY--- 785
DB 666 LKIKFLKAGIYEPPIITTSGNPKSNISILRVKVCQCSNBDP-----TDVDRIVGA 718
QY 786 GMAVGILLFTLIVIGIILAVFIRIKDKDKDNVESQAQSEVAP 829
DB 719 GLCTGATIALILCTIILILIVLMFVVMKRRDKERQAKQLLIDP 762

RESULT 13
AAW25658
ID AAW25658 standard; Protein; 916 AA.
XX
AC AAW25658;
XX
DT 04-NOV-1997 (first entry)
XX
DE Human cadherin-4.
XX
KW Human; cadherin; rat; calcium-dependent cell adhesion protein;
KM superfamily; cytoskeleton; eatenin; cancer.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Msc-difference 78 /note= "Encoded by AAA"
FT
XX
PN US5646250-A.
PD 08-JUL-1997.
XX
PF 17-APR-1992; 92US-0872643.
XX
PR 19-APR-1993; 93US-0049460.
PR 17-APR-1992; 92US-0872643.
PR 01-NOV-1994; 94US-0332638.
XX
PA (DOHE-) DOHENY EYE INST.
PI Suzuki S;
PI
XX
DR WPI; 1997-362997/33.
DR N-PSDB; AAT85433.
XX
PT Human and rat cadherin polypeptide(s) - mediate cell-cell adhesion
PT
XX
PS Example 2; Column 61-66; 56pp; English.
XX
XX This sequence represents human cadherin-4. The invention specifically
CC provides details of human cadherin-5, -8, -11, -12 and -13, and rat
CC cadherin-8, -5, -11 and -13. Cadherins are calcium-dependent cell
CC adhesion proteins. They are glycosylated integral membrane proteins
CC that have an N-terminal extracellular domain that determines binding
CC specificity, a hydrophobic membrane spanning region and a C-terminal
CC cytoplasmic domain, which is highly conserved among members of the
CC superfamily. The C-terminal domain interacts with the cytoskeleton
CC through eatenin and other cytoskeleton-associated proteins. The
CC novel cadherin proteins may be used in the analysis of the role of
CC cadherins in various cancers. Sequence analysis of the cadherin
CC proteins also allows investigation of the structure and function of
CC cadherin. The cadherin proteins may be isolated by using anti-cadherin
CC antibodies. These antibodies may also be used to modulate the activity
CC of cadherin and to determine the tissue specific distribution of cadherin

```

CC proteins. Each subclass of cadherins has a unique tissue distribution
 CC pattern.

XX Sequence 916 AA;

Query Match 15.8%; Score 684.5; DB 18; Length 916;
 Best Local Similarity 29.3%; Pred. No. 2.3e-46;
 Matches 176; Conservative 106; Mismatches 288; Indels 31; Gaps 12;

239 WKAPKPEVMEVNTDHPRIKITOVW---NDPCAQVSL--VDKELPRFPESID-QEGDI 292
 171 WVIP-PINVEBNSGPFPOQLVIRSDKNDIPRISITGVADQPMVEFINSMSGRM 229
 293 YVTPQLDREKDAYVFAVAKDEYKPLSYPLEIHVKVDINDNPPTCPSPVTFVQEN 352
 230 YVTRPMDREHASYHLRAHAVDMNGKVENPIDIYIVIDMNDHPEFINQVNCSDG 289
 353 ERLGNSIGTLTAHDEENTANSFLNRYIVEQTPKLPMDGLFLIQYAGMLQAKOSLKK 412
 290 SKPGTYMTITANDADSTTANGMVRVRYIYQTPQSPSQMFTINSBTGDIVTVAAGMDR 349
 413 QDPFQVNLITIEVSDK-----FKTLCEVOINVINDIQPIPEKSDYGNLTLAEDTNG 466
 350 EKVOQYTVIVQATDMEGNLYGLSNTATITVTDVNDNPSFTASTFAG--EVPENSV 407
 467 STILTIQATDADEPFTGSSKILYHIKDSBGRGVDTPHTNTGYVILKKPLDFETAAV 526
 408 TVVANLTVMRDQPHSPNNMAVRYIISGDPGSHSVATDPTVMGQWTVVKADYELNRA 467
 527 SNIVFKAENPEPLVFGVKNASFAKFTLVTDVNEAPQSOHVPOAKSEDAVIGTKVG 586
 468 FMLTVWVNSQAPLASIGMSFOSTAGVTISIMDINEAPYPSNHKILRLEGVPPTGLT 527
 587 NTAKDPEGL---DISYSLRGDTFGMLKIDHVTGEISVAPLDREA---GSPYRQVVA 639
 528 TFSAVDDPRMOQAVRYSKLSDPASWLNATNGQITVAVLRESLYTKNNYEATFLA 587
 640 TEVGSGLSSVSEFHLILMDVNDNPRRLADYTGJFCHPLASAGSLIFEATDDQHLFR 639
 588 ADNGIPPASGTGLQIYLIDINDAPRLPKRAQICERPNLNA---INTTAADADVHNI 644
 700 GPH-FTFSLGSGSLQNDMEVSKINGTHARLSTRHTDFEERAVVYLIRINDGRRPLEGIV 758
 645 GPYVFELPFVPAVRKMWITTRLNGDYAQLSLRILYLEAGMDVPIIVTDSGNPPLSNTS 704
 759 SLPTFGSCVSGSCFRPAGHGTGIPVGMVAGILTLTLVIGIILAV--FIRIKDKK 816
 705 IIKVKVCPDNDGDCCTTIG--AVAAAGLGTGAIVAILICILITLTVLFLVMMKRREK 761
 QY 817 D 817
 Db 762 E 762
 RESULT 14
 AAM13129
 ID AAM13129 standard; Protein; 916 AA.
 XX AAM13129;
 AC AAM13129;
 DT 13-MAY-1997 (first entry)
 XX Full length human cadherin-4.
 DE Ca2+ dependent; cell adhesion protein; foetal; cadherin; rat;
 KW brain; human; antibody; purification; determination;
 XX tissue expression; binding antagonist; calcium ion.
 OS Homo sapiens.
 XX US559725-A.
 PN 28-JUN-1997.
 PD

XX 17-APR-1992; 92US-0872643.
 PF 26-JUN-1994; 94US-0188228.
 PR 17-APR-1992; 92US-0872643.
 PR 19-APR-1993; 93US-0049460.
 XX (DOHE-) DOHENY EYE INST.
 PI Suzuki S;
 XX WPI, 1997-108328/10.
 DR N-PSDB; AAT61920.
 XX Antibodies to cadherin proteins - useful as cadherin antagonists,
 PT etc.
 PS Claim 1; Columns 67-72; 59pp; English.
 CC The present sequence is full length human cadherin-4, which
 CC is a Ca2+ dependent cell adhesion protein. The human cadherin cDNA
 CC was isolated from a foetal brain cDNA library, using probes based
 CC on homologous rat cadherin cDNA.
 CC Antibodies or fragments that specifically bind the human cadherin
 CC can be used to purify the cadherin, determine its tissue expression
 CC and antagonise its ligand/antiligand binding activities.
 CC
 SQ Sequence 916 AA;
 Query Match 15.8%; Score 684.5; DB 18; Length 916;
 Best Local Similarity 29.3%; Pred. No. 2.3e-46;
 Matches 176; Conservative 106; Mismatches 288; Indels 31; Gaps 12;

239 WKAPKPEVMEVNTDHPRIKITOVW---NDPCAQVSL--VDKELPRFPESID-QEGDI 292
 171 WVIP-PINVEBNSGPFPOQLVIRSDKNDIPRISITGVADQPMVEFINSMSGRM 229
 293 YVTPQLDREKDAYVFAVAKDEYKPLSYPLEIHVKVDINDNPPTCPSPVTFVQEN 352
 230 YVTRPMDREHASYHLRAHAVDMNGKVENPIDIYIVIDMNDHPEFINQVNCSDG 289
 353 ERLGNSIGTLTAHDEENTANSFLNRYIVEQTPKLPMDGLFLIQYAGMLQAKOSLKK 412
 290 SKPGTYMTITANDADSTTANGMVRVRYIYQTPQSPSQMFTINSBTGDIVTVAAGMDR 349
 413 QDPFQVNLITIEVSDK-----FKTLCEVOINVINDIQPIPEKSDYGNLTLAEDTNG 466
 350 EKVOQYTVIVQATDMEGNLYGLSNTATITVTDVNDNPSFTASTFAG--EVPENSV 407
 467 STILTIQATDADEPFTGSSKILYHIKDSBGRGVDTPHTNTGYVILKKPLDFETAAV 526
 408 TVVANLTVMRDQPHSPNNMAVRYIISGDPGSHSVATDPTVMGQWTVVKADYELNRA 467
 527 SNIVFKAENPEPLVFGVKNASFAKFTLVTDVNEAPQSOHVPOAKSEDAVIGTKVG 586
 468 FMLTVWVNSQAPLASIGMSFOSTAGVTISIMDINEAPYPSNHKILRLEGVPPTGLT 527
 587 NTAKDPEGL---DISYSLRGDTFGMLKIDHVTGEISVAPLDREA---GSPYRQVVA 639
 528 TFSAVDDPRMOQAVRYSKLSDPASWLNATNGQITVAVLRESLYTKNNYEATFLA 587
 640 TEVGSGLSSVSEFHLILMDVNDNPRRLADYTGJFCHPLASAGSLIFEATDDQHLFR 639
 588 ADNGIPPASGTGLQIYLIDINDAPRLPKRAQICERPNLNA---INTTAADADVHNI 644
 700 GPH-FTFSLGSGSLQNDMEVSKINGTHARLSTRHTDFEERAVVYLIRINDGRRPLEGIV 758
 645 GPYVFELPFVPAVRKMWITTRLNGDYAQLSLRILYLEAGMDVPIIVTDSGNPPLSNTS 704
 759 SLPTFGSCVSGSCFRPAGHGTGIPVGMVAGILTLTLVIGIILAV--FIRIKDKK 816
 705 IIKVKVCPDNDGDCCTTIG--AVAAAGLGTGAIVAILICILITLTVLFLVMMKRREK 761

QY 817 D 817
Db 762 E 762

RESULT 15

AAR63533 standard; Protein; 896 AA.

AA63533;

07-JUN-1995 (first entry)

Human HT-1376 cell-derived Leukocyte Chemotactic Factor.

rHT-LCF; HT-1376 cell derived leukocyte chemotactic factor;

immunopotentiator; monocyte and macrophage migration factor;

wound healing.

Homo sapiens (cell line HT-1376).

Key Location/Qualifiers

FT Peptide

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

QY 409 SLKKDPTQYNLTIEVSDKPF-----TLCEVOINVINDIQPIPEKSDYGNLTAE 461

Db 310 YLDREAVDXYSLIMKQMDGCPFGILIGTSTCI--ITVDSNDNAFTFQNAE--AFVE 365

QY 462 DTNIGSTIILTIQATDADEPFTGSSKILYHIIKGDSEGLGVDTPPTNGYVILKKPLDF 521

Db 366 ENAFNVEILRIPIEDDLINTANMRVNFILKGNENGHKISTDKETNGVLSVAKPLAN 425

QY 522 ETAAVSNIVFKAEPEPLVFGV--KINASSPAKFTLIYTVNNEAPQSQHVQAKVSEDA 580

Db 426 EENRQVNLBIGNVNEAPFARDIPRVYALNRALVTVAHRDLDEGEPECTPAQVVRIKENLA 485

QY 581 IGTKGVNVTAKDPE--GLDISLSLGDTRGMKIDHVTGEIJSVAPLDEAGSP----Y 633

Db 486 VSKINGYKAYDENENGLRYKHLDRGMATTIDEIGSIIITSKILDBREVTFRQNELY 545

QY 634 RYQVAVTEVGSLSVSEFHLIMDVNDNPPRLAKDYGLFPCHPLSAPGSLIFEATD 693

Db 546 NITVLAIDKDDRECTGTAVN--IEDVNDNPEIIGEX--VVICRPMGYTDIL--AVDP 599

QY 694 DQHLFRGPHFTFSLGSGS--LQNDMEVSKINGTHARLS--TRHTDEERAYVILIRINDG 750

Db 600 DEPVHGAP--FYFSLPMTSPSEISRLMSLTQVNDPAELSYQKAGFQE--YTPITVKD-- 654

QY 751 RPLEGIVSLPYTFGSCVBSGCPRPAGHQGIPVGMAY--GIILTLTVIGIILAVFIR 809

Db 655 RAGQAAIKLRVNLCECTHTPQCRATSRSTGVLGKMAILIILGIALLSFVLLTVCGV 714

QY 810 IKKDKG 816

Db 715 FGATKGR 721

Search completed: June 22, 2003, 01:04:37
Job time : 66 secs

Monocyte and macrophage migration factor protein - is
immunopotentiator and vulnerary and is isolated from culture of
HT-1376 cell line by recombinant methods

Claim 5; Page 34-41; 68pp; Japanese.

Sequences coding for proteins having monocyte and macrophage
migration factor activity were isolated from a cDNA library
prepared by the isolated cDNA clones are useful as
immunopotentiators for treatment of infections, immune
deficiency diseases and cancer. The proteins are also useful in
wound healing preparations and for treatment of skin diseases.

Sequence 896 AA;

Query Match 15.5%; Score 671.5; DB 15; Length 896;
Best Local Similarity 30.8%; Pred. No. 2.5e-45;
Matches 187; Conservative 106; Mismatches 263; Indels 51; Gaps 22;

QY 239 WKAPKPVEMVENSIDPHIKITQVRMNDGAGYSL-----VDKEKLPFPFSSIDQEG 290

Db 137 W-APICSMQENSIGPFLPLQVE--SPAQNTVTFYISISGRGVKPEPLNLFYIERD-TG 193

QY 291 DIVYTOPLDREKAY--VEYAVAKDEXGKPLSYPLHIVKDIINDNPPRCSPVTVPE 348

Db 194 NLPCTRPVDRREYVFDLIAASTADGYSADL--PLPLPIRVEDENDNHPVFTETATYNFE 251

QY 349 VOENNERLGNISIGTLTAHDEENTANSFLNRIYEQTPKLPMDGLFLIQTYAGMLQAKQ 408

Db 252 VLSSRPQTTGVGCADTRDEPDMTRKLSILQQTFRSP--GLFSVHPSTGVITTVSH 309

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2003, 01:03:29 ; Search time 35 seconds
(without alignments)
699.425 Million cell updates/sec

Title: US-10-025-380-1081
Perfect score: 4321
Sequence: 1 MIQAHLSICLMLTYLATG.....DKGKNVESQAQSEVPELRS 832

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4321	100.0	832	1 US-08-431-560-1	Sequence 1, Appl1
2	4321	100.0	832	1 US-08-463-345-1	Sequence 1, Appl1
3	696	16.1	913	2 US-08-474-067-6	Sequence 6, Appl1
4	696	16.1	913	2 US-08-474-068A-6	Sequence 6, Appl1
5	696	16.1	913	2 US-08-472-481-5	Sequence 5, Appl1
6	684.5	15.8	916	1 US-08-188-228-48	Sequence 46, Appl1
7	684.5	15.8	916	1 US-08-332-643-42	Sequence 42, Appl1
8	684.5	15.8	916	1 US-08-332-638-48	Sequence 48, Appl1
9	657.5	15.2	713	1 US-08-188-228-62	Sequence 62, Appl1
10	657.5	15.2	713	1 US-08-332-643-56	Sequence 56, Appl1
11	657.5	15.2	713	1 US-08-332-638-62	Sequence 62, Appl1
12	650.5	15.1	555	2 US-08-453-702A-98	Sequence 98, Appl1
13	634	14.7	837	2 US-08-474-067-7	Sequence 7, Appl1
14	634	14.7	837	2 US-08-474-068A-7	Sequence 7, Appl1
15	634	14.7	837	2 US-08-472-481-6	Sequence 6, Appl1
16	633	14.6	556	1 US-07-998-003A-98	Sequence 98, Appl1
17	633	14.6	556	1 US-08-453-774B-98	Sequence 98, Appl1
18	633	14.6	556	1 US-08-453-695A-98	Sequence 98, Appl1
19	633	14.6	556	1 US-08-268-161A-98	Sequence 98, Appl1
20	633	14.6	556	1 US-09-099-639-98	Sequence 98, Appl1
21	630.5	14.6	712	2 US-08-474-067-2	Sequence 2, Appl1
22	630.5	14.6	712	2 US-08-474-067-5	Sequence 5, Appl1
23	630.5	14.6	712	2 US-08-474-068A-2	Sequence 2, Appl1
24	630.5	14.6	712	2 US-08-474-068A-5	Sequence 5, Appl1
25	630.5	14.6	712	2 US-08-472-481-2	Sequence 2, Appl1
26	630.5	14.6	717	2 US-08-474-067-4	Sequence 4, Appl1
27	630.5	14.6	717	2 US-08-474-068A-4	Sequence 4, Appl1

28	630.5	14.6	717	2 US-08-472-481-4	Sequence 4, Appl1
29	608	14.1	884	2 US-08-474-067-8	Sequence 8, Appl1
30	608	14.1	884	2 US-08-474-068A-8	Sequence 8, Appl1
31	608	14.1	884	2 US-08-472-481-7	Sequence 7, Appl1
32	593	13.7	878	1 US-08-237-919-2	Sequence 2, Appl1
33	593	13.7	878	1 US-08-732-429-2	Sequence 2, Appl1
34	593	13.7	878	4 US-09-798-267-3	Sequence 3, Appl1
35	593	13.7	878	4 US-09-798-267-3	Sequence 3, Appl1
36	593	13.7	878	4 US-09-798-267-3	Sequence 3, Appl1
37	592.5	13.7	117	4 US-09-187-859-25	Sequence 25, Appl1
38	574.5	13.3	796	2 US-08-738-349-2	Sequence 2, Appl1
39	560.5	13.0	796	2 US-08-738-349-4	Sequence 4, Appl1
40	559	12.9	822	2 US-08-474-067-9	Sequence 9, Appl1
41	559	12.9	822	2 US-08-474-068A-9	Sequence 9, Appl1
42	559	12.9	822	2 US-08-472-481-8	Sequence 8, Appl1
43	555	12.8	106	1 US-08-326-117B-6	Sequence 6, Appl1
44	555	12.8	106	3 US-08-982-129-6	Sequence 6, Appl1
45	554	12.8	106	4 US-09-178-176B-11	Sequence 11, Appl1

ALIGNMENTS

```
RESULT 1
US-08-431-560-1
; Sequence 1, Application US/08431560
; Patent No. 5620855
; GENERAL INFORMATION:
; APPLICANT: Anne H. Danczig, et al.
; TITLE OF INVENTION: Mammalian Influx Peptide
; TITLE OF INVENTION: Transporter
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,560
; FILING DATE: 01-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/013,462
; FILING DATE: 04-FEB-1993
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 832 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-431-560-1

Query Match 100.0%; Score 4321; DB 1; Length 832;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 832; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIQAHLSICLMLTYLATGYGQEGKFSGLPKMTFSIYEGQSPQITFOFKANPPAVTF 60
1 MIQAHLSICLMLTYLATGYGQEGKFSGLPKMTFSIYEGQSPQITFOFKANPPAVTF 60
Db 1 MIQAHLSICLMLTYLATGYGQEGKFSGLPKMTFSIYEGQSPQITFOFKANPPAVTF 60
QY 61 ELTGERDNI FVIERREGLLYYNRALDRETSSTNHLQVAAIDANGIYEGVPITIEVKDIN 120
61 ELTGERDNI FVIERREGLLYYNRALDRETSSTNHLQVAAIDANGIYEGVPITIEVKDIN 120
Db 61 ELTGERDNI FVIERREGLLYYNRALDRETSSTNHLQVAAIDANGIYEGVPITIEVKDIN 120
QY 121 DNRPTFLQSKYEGSVQRNRPKPLYYNATDLDDEPATNGQLYYQIVQLPNNINMYTF 180
121 DNRPTFLQSKYEGSVQRNRPKPLYYNATDLDDEPATNGQLYYQIVQLPNNINMYTF 180
```

Db 121 DNRPTFLOSKYEGSVQRNSRPGKPEFLVYNATLDDPATPNQLYQIVIQLEPMINMYF 180
Qy 181 QINNKTAISLTRGSOELNPAKPSYNLVISYKMGQSGNSFSDTSDVIITENIMK 240
Db 181 QINNKTAISLTRGSOELNPAKPSYNLVISYKMGQSGNSFSDTSDVIITENIMK 240
Qy 241 APRVENVENSTDPHPKIKITQVRNMDPGAQYSLVDKELPRFPFSIDQEGDIYVTOPLDR 300
Db 241 APRVENVENSTDPHPKIKITQVRNMDPGAQYSLVDKELPRFPFSIDQEGDIYVTOPLDR 300
Qy 301 EKKDAYFYVAADKDEYKPLSYPLEIHVKYKINDNPTCPSPYTVFEVQENRGLNSIG 360
Db 301 EKKDAYFYVAADKDEYKPLSYPLEIHVKYKINDNPTCPSPYTVFEVQENRGLNSIG 360
Qy 361 TLTAHDBDEENTANSFLNRYIVEQTPKLPMDGLFIQTYAGMLQAKOSLKKODTPQYNL 420
Db 361 TLTAHDBDEENTANSFLNRYIVEQTPKLPMDGLFIQTYAGMLQAKOSLKKODTPQYNL 420
Qy 421 TIEVSDKDFKTLCEVQINVIDINDQIPIFEKSDYGNLTAEADTNGSTILTITQATDADEP 480
Db 421 TIEVSDKDFKTLCEVQINVIDINDQIPIFEKSDYGNLTAEADTNGSTILTITQATDADEP 480
Qy 481 FTGSSKILYHIHIIKGDSEGRIGVDTDPHTNTGYVIIKKPLDEFTAAVSNIVFKANPEPLV 540
Db 481 FTGSSKILYHIHIIKGDSEGRIGVDTDPHTNTGYVIIKKPLDEFTAAVSNIVFKANPEPLV 540
Qy 541 FGKYNASSFAKFTLIYTDVNEAPOSQHVFOAKVSEDAVIGTKVGNVTAODEGLDISY 600
Db 541 FGKYNASSFAKFTLIYTDVNEAPOSQHVFOAKVSEDAVIGTKVGNVTAODEGLDISY 600
Qy 601 SLRGDTGMLKIDHVTGEIFSVAPLDRAGSPYRVQVATEVGGSSLSVSEFHLILMDV 660
Db 601 SLRGDTGMLKIDHVTGEIFSVAPLDRAGSPYRVQVATEVGGSSLSVSEFHLILMDV 660
Qy 661 NDNPRLAKDYTGIFCHPLSAPGSLIFEATDDQHLFRGPHFTFSIGSSGLQNDMEVSK 720
Db 661 NDNPRLAKDYTGIFCHPLSAPGSLIFEATDDQHLFRGPHFTFSIGSSGLQNDMEVSK 720
Qy 721 INGHARLSTRHTDPEERAYVVLIRINDGRPPLEGIVSLRVTCSCVEGSCFRPAGHOT 780
Db 721 INGHARLSTRHTDPEERAYVVLIRINDGRPPLEGIVSLRVTCSCVEGSCFRPAGHOT 780
Qy 781 GIPVGMAGVILITLTLVIGIILAVVFIKIKDKGKNVESAQASEVKKPLRS 832
Db 781 GIPVGMAGVILITLTLVIGIILAVVFIKIKDKGKNVESAQASEVKKPLRS 832

RESULT 2
US-08-463-345-1
Sequence 1, Application US/08463345
Patent No. 5710018
GENERAL INFORMATION:
APPLICANT: Anne H. Dantzig, et al.
TITLE OF INVENTION: Mammalian Influx Peptide
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN
COUNTRY: U.S.A.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,345
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/013,462
FILING DATE: 04-FEB-1993
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-345-1

Query Match 100.0%; Score 4321; DB 1; Length 832;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 832; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MILQAHLSCLMLATATGYGEGKFSGLKMTFSIYEGQPSQIIFQFKANPAVTF 60
Db 1 MILQAHLSCLMLATATGYGEGKFSGLKMTFSIYEGQPSQIIFQFKANPAVTF 60
Qy 61 ELTGEDNIFVIEREGILYYNRALDBETRSTNMLQVAALDANGIIVEGPVITIEVKDIN 120
Db 61 ELTGEDNIFVIEREGILYYNRALDBETRSTNMLQVAALDANGIIVEGPVITIEVKDIN 120
Qy 121 DNRPTFLOSKYEGSVQRNSRPGKPEFLVYNATLDDPATPNQLYQIVIQLEPMINMYF 180
Db 121 DNRPTFLOSKYEGSVQRNSRPGKPEFLVYNATLDDPATPNQLYQIVIQLEPMINMYF 180
Qy 181 QINNKTAISLTRGSOELNPAKPSYNLVISYKMGQSGNSFSDTSDVIITENIMK 240
Db 181 QINNKTAISLTRGSOELNPAKPSYNLVISYKMGQSGNSFSDTSDVIITENIMK 240
Qy 241 APRVENVENSTDPHPKIKITQVRNMDPGAQYSLVDKELPRFPFSIDQEGDIYVTOPLDR 300
Db 241 APRVENVENSTDPHPKIKITQVRNMDPGAQYSLVDKELPRFPFSIDQEGDIYVTOPLDR 300
Qy 301 EKKDAYFYVAADKDEYKPLSYPLEIHVKYKINDNPTCPSPYTVFEVQENRGLNSIG 360
Db 301 EKKDAYFYVAADKDEYKPLSYPLEIHVKYKINDNPTCPSPYTVFEVQENRGLNSIG 360
Qy 361 TLTAHDBDEENTANSFLNRYIVEQTPKLPMDGLFIQTYAGMLQAKOSLKKODTPQYNL 420
Db 361 TLTAHDBDEENTANSFLNRYIVEQTPKLPMDGLFIQTYAGMLQAKOSLKKODTPQYNL 420
Qy 421 TIEVSDKDFKTLCEVQINVIDINDQIPIFEKSDYGNLTAEADTNGSTILTITQATDADEP 480
Db 421 TIEVSDKDFKTLCEVQINVIDINDQIPIFEKSDYGNLTAEADTNGSTILTITQATDADEP 480
Qy 481 FTGSSKILYHIHIIKGDSEGRIGVDTDPHTNTGYVIIKKPLDEFTAAVSNIVFKANPEPLV 540
Db 481 FTGSSKILYHIHIIKGDSEGRIGVDTDPHTNTGYVIIKKPLDEFTAAVSNIVFKANPEPLV 540
Qy 541 FGKYNASSFAKFTLIYTDVNEAPOSQHVFOAKVSEDAVIGTKVGNVTAODEGLDISY 600
Db 541 FGKYNASSFAKFTLIYTDVNEAPOSQHVFOAKVSEDAVIGTKVGNVTAODEGLDISY 600
Qy 601 SLRGDTGMLKIDHVTGEIFSVAPLDRAGSPYRVQVATEVGGSSLSVSEFHLILMDV 660
Db 601 SLRGDTGMLKIDHVTGEIFSVAPLDRAGSPYRVQVATEVGGSSLSVSEFHLILMDV 660
Qy 661 NDNPRLAKDYTGIFCHPLSAPGSLIFEATDDQHLFRGPHFTFSIGSSGLQNDMEVSK 720
Db 661 NDNPRLAKDYTGIFCHPLSAPGSLIFEATDDQHLFRGPHFTFSIGSSGLQNDMEVSK 720
Qy 721 INGHARLSTRHTDPEERAYVVLIRINDGRPPLEGIVSLRVTCSCVEGSCFRPAGHOT 780
Db 721 INGHARLSTRHTDPEERAYVVLIRINDGRPPLEGIVSLRVTCSCVEGSCFRPAGHOT 780
Qy 781 GIPVGMAGVILITLTLVIGIILAVVFIKIKDKGKNVESAQASEVKKPLRS 832
Db 781 GIPVGMAGVILITLTLVIGIILAVVFIKIKDKGKNVESAQASEVKKPLRS 832

RESULT 3
US-08-474-067-6

Sequence 6, Application US/08474067
Patent No. 581518
GENERAL INFORMATION:
APPLICANT: Ranscht, Barbara
TITLE OF INVENTION: T-Cadherin Adhesion Molecule
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,067
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/213,361
FILING DATE: 14-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/607,293
FILING DATE: 30-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1682
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 913 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-474-067-6

Query Match 16.1%; Score 696; DB 2; Length 913;
Best Local Similarity 26.9%; Pred. No. 6.9e-53;
Matches 209; Conservative 117; Mismatches 328; Indels 124; Gaps 22;
QY 126 FLOSKYGSVRQNSRPEKPFYLYNATLDDPATNGQLTYQIYQLPMINNMATFOIN-- 183
DB 41 FPDVHSAVVSRSVHGQPLNVRFGSCDE-----NRKIYFGSSEP 81
QY 184 -----NKTGALISLTREGSQELNPAKNSYMLVISVKDMGQS-----ENSPGDTT 228
DB 82 EDRFVGEVDGVVYAKRSQLSAEPE-----FVVSARKETOEKQMKVTKLTPPATFGAS 136
QY 229 SVDIIVTENT-----WKAPKPVEMVNSTDPHPIKITOVWRMDPGA 269
DB 137 EKQOKKIEDIIFPMQYKQSSHLKQKQDMVIFP-PINLPENSGPFFQELVRR----- 169
QY 270 QYSLVDEKELPRPFESL-----DQ-----EGDIYVTOPLDREEDAVFYAVAK 313
DB 190 -----SPDDKSLSRYSVTGPGADQPTGIFINPISQSLSVTKPLDEQQLASPHLAHAV 245
QY 314 DEYKPLSYPLLEIHVKQDINDNPPCTGSPVTVFEVQENRSLGNSIGTLTAHDESENTA 373
DB 246 DVANGQVENIDIVINIDANDNRPEFLHVMGTVFEGSKPFGYVWTVTAIADDPNAQ 305
QY 374 NSFLNTRIYEQTPLPMDGLFLIQTAYAGMLQALAKQSLKQDTPQYMLTIEVSQD----- 428
DB 306 NGMLRYRILISQASSPSPNMTFINNETGDIITVAAGLDREKVOQYTLIIQATMEGNPTV 365
QY 429 -FTLCTVQVQVNVNDINDQIPF-EKSDYKMLTAEEDINISITLITATQATADDPFGSSK 486
DB 366 GLSNATATVITVDVNDNPEFTAMTFYGEV---PENRVDVIVANLITVDKQDPHTPAMN 422

QY 487 ILVHIKGSSEGLGVDTPDHTNTGYVILIKKPLDFETAASNVFAENBEPVFGVKYN 546
DB 423 ARYQMGGDPTGQFTLLTDPNSNDGLVTVVKPIDFENNRHFVLTVAENQVPLAKGIQHP 482
QY 547 ASSFAKTLIVTVNNEAPQSGHVFOAKSDEVALTGKGNVTAQD-----EGLDISYSL 602
DB 483 POSTATVSIIVDVNSFPVFPNPKLVROEBGLASMLTTFARPDRTYMOQTSIRYSK 542
QY 603 RGTGRLKIDHTVGLFSAFLDREA-----GSPYVQVATVGGSSLSVSEFHLIM 658
DB 543 LSDPAMKIDPVNGQITTTAVLDRESIVQNMNMTATFLASDNGIIPMSGTGTQIYLL 602
QY 659 DVANDNPPRL-AKDYTGFLFCHPLSAPGSLFEATDDQHLFRGPHFTFSISGQ--SHQND 715
DB 603 DINDMPOVNPKRAT---TCEITQ-ENAINITAVDDIDPNAGP-PAFELPDSPPSIRKN 657
QY 716 WEVSKINGTHARLSTHTDEERAYVVLIRINDGAPPLEGIVSLPVTFCSC-VEGSCFR 774
DB 658 WTVIRISGDHAQSLRIFLEAGIYDVPYITDSGNPHASSTSVLKVKVCQCDINGDC-- 715
QY 775 PACHQGIPTV---GNAVGILLTLVLVIGIILAVFIRIKKQKXNVESAQASVYKP 829
DB 716 -----TDVDRIVAGLGTGAIITAILCITILLVLIMFVMMKRRDKERQAKQLIDP 768
RESULT 4
US-08-474-068A-6
Sequence 6, Application US/08474068A
Patent No. 5837525
GENERAL INFORMATION:
APPLICANT: Ranscht, Barbara
TITLE OF INVENTION: T-Cadherin Adhesion Molecule
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,068A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/213,361
FILING DATE: 14-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/607,293
FILING DATE: 30-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1683
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 913 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-474-068A-6
Query Match 16.1%; Score 696; DB 2; Length 913;
Best Local Similarity 26.9%; Pred. No. 6.9e-53;
Matches 209; Conservative 117; Mismatches 328; Indels 124; Gaps 22;

```

QY 126 FLOSKYEGSVRONSREKPEFLVYNATLDDPATNGOLYYQIVQLPMINWYFOIN-- 183
DB 41 FPEVDHSAVVSRSVHGQPLNVRFGSCDE-----NRKIYFGSSSEP 81
QY 184 -----NKTGAISLTREGSGEINPAKNPSYNLVSVKMGQS-----ENSFSDTJ 228
DB 82 EDRVGEDGVYAERSFQLSAEPT-----FVVSARKETQEBQMKVKTLPBPAFTGAS 136
QY 229 SVDIIVTENT-----WKAPKPVEMVENSIDPPHPIKITQVRMNDPGA 269
DB 137 EKQOKKIEDIIFPMQYKXSSHLKROKRDWVIR-PINLPENSRRPFGQELVRIR----- 189
QY 270 QYSLVDEKXLPREFSI-----DQ-----EGDIYVTOPLDREKDAVYFVAVAK 313
DB 190 -----SDRDKSLSLRYSVTGPGADQPPGTGIFINPISQSLSTKPLDREQIASPHLRHAHV 245
QY 314 DEYKPLSYPLEIHVKAKDINDNPTCPSPVTVPEVOENBERLGNISGTLTAHOREENTA 373
DB 246 DVNGNOVENPIDIVINIDMNDNRPFLHQVWNGTVEGSKPGTYVMTVTAIDADDPNAQ 305
QY 374 NSFLNRYIVEQTPKLPMDGLFLIQTVAQMLQAKOSLKKODTPQYNLTIEVSKD----- 428
DB 306 NGMLRYRILSQAFSSPSPMNETINNTEGDIITVAAGLDREKVOQYTLIIQATMEGNPTY 365
QY 429 -FKTLCFVQINVIDINDQIPF-EKSDYGNLTLAEDTNIGSTLITQATDADEPFTGSSK 486
DB 366 GLSNTATAVITVTDVNDNPEFTAMTFYGEV---PENRVDIVIANLTVTDKQDPTPAMN 422
QY 487 ILVHIHIGDSEGRGLGVDTPHTNTGYIILKKPLDFTAAVSNIVFKAENPEPLVFGVKN 546
DB 423 ARYOMTGGDPPTGQITLIDPNSNGLTAVKPIDIFETNRMPVLTVAAENOVPLAKGIQHP 482
QY 547 ASSFAKFTLLVTDVNEAPQPSQHVFOAKVSEDAVIGTKVGNVTAKP---EGLDISYSL 602
DB 483 POSTATVTSIVTDVNESPYFVNPVKLVQEBGLAGSMLTFTFARDDRYVQOTSLSYK 542
QY 603 RGDTRGMLKIDHVTGEIFSVAPLDRFA-----GSPRYQVAVTEVSGSSLSVSFHLILM 658
DB 543 LSDPANMLKIDPVNGQITTTAVLDRRESIYQNNMNYATFLASDNGIPMSGTGLQIYLL 602
QY 659 DVNDNPPRL-AKDYTGFLFCHPLSAPGSLIFEATDDOHLFRGPHFTFSLSG--SLQND 715
DB 603 DINNAQVNPKEAT---TCETLQ-PNAINITAVDPDIDPNAQ-PAFELPDSPPSIRKN 657
QY 716 WEYSKINGTARLSTRHTDEERAYVVLIRINDGRRPPLBGIVSLPVTFGSC-VEGSCFR 774
DB 658 WTIVRISGDHQLSLRIRFLBAGIYDVPVITDSGNPFIASSTVLKXVCCOCDINGDC-- 715
QY 775 PAGOQTOPIPV---GMAVGLITLLVIGILLAVFIRIKKDKKDNVESQASEVRP 829
DB 716 ----TVDVIRVAGLGTGAILITLITLILVLMFVMMKRRDKEROAKQILLDP 768

RESULT 5
US-08-472-481-5
; Sequence 5, Application US/08472481
; Patent No. 5863804
; GENERAL INFORMATION:
; APPLICANT: Ranscht, Barbara
; TITLE OF INVENTION: T-Cadherin Adhesion Molecule
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,481
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/213,361
; FILING DATE: 14-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/607,293
; FILING DATE: 30-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1686
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 913 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-472-481-5

Query Match 16.1%; Score 696; DB 2; Length 913;
Best Local Similarity 26.9%; Pred. No. 6,9e-53;
Matches 209; Conservative 117; Mismatches 328; Indels 124; Gaps 22;

QY 126 FLOSKYEGSVRONSREKPEFLVYNATLDDPATNGOLYYQIVQLPMINWYFOIN-- 183
DB 41 FPEVDHSAVVSRSVHGQPLNVRFGSCDE-----NRKIYFGSSSEP 81
QY 184 -----NKTGAISLTREGSGEINPAKNPSYNLVSVKMGQS-----ENSFSDTJ 228
DB 82 EDRVGEDGVYAERSFQLSAEPT-----FVVSARKETQEBQMKVKTLPBPAFTGAS 136
QY 229 SVDIIVTENT-----WKAPKPVEMVENSIDPPHPIKITQVRMNDPGA 269
DB 137 EKQOKKIEDIIFPMQYKXSSHLKROKRDWVIR-PINLPENSRRPFGQELVRIR----- 189
QY 270 QYSLVDEKXLPREFSI-----DQ-----EGDIYVTOPLDREKDAVYFVAVAK 313
DB 190 -----SDRDKSLSLRYSVTGPGADQPPGTGIFINPISQSLSTKPLDREQIASPHLRHAHV 245
QY 314 DEYKPLSYPLEIHVKAKDINDNPTCPSPVTVPEVOENBERLGNISGTLTAHOREENTA 373
DB 246 DVNGNOVENPIDIVINIDMNDNRPFLHQVWNGTVEGSKPGTYVMTVTAIDADDPNAQ 305
QY 374 NSFLNRYIVEQTPKLPMDGLFLIQTVAQMLQAKOSLKKODTPQYNLTIEVSKD----- 428
DB 306 NGMLRYRILSQAFSSPSPMNETINNTEGDIITVAAGLDREKVOQYTLIIQATMEGNPTY 365
QY 429 -FKTLCFVQINVIDINDQIPF-EKSDYGNLTLAEDTNIGSTLITQATDADEPFTGSSK 486
DB 366 GLSNTATAVITVTDVNDNPEFTAMTFYGEV---PENRVDIVIANLTVTDKQDPTPAMN 422
QY 487 ILVHIHIGDSEGRGLGVDTPHTNTGYIILKKPLDFTAAVSNIVFKAENPEPLVFGVKN 546
DB 423 ARYOMTGGDPPTGQITLIDPNSNGLTAVKPIDIFETNRMPVLTVAAENOVPLAKGIQHP 482
QY 547 ASSFAKFTLLVTDVNEAPQPSQHVFOAKVSEDAVIGTKVGNVTAKP---EGLDISYSL 602
DB 483 POSTATVTSIVTDVNESPYFVNPVKLVQEBGLAGSMLTFTFARDDRYVQOTSLSYK 542
QY 603 RGDTRGMLKIDHVTGEIFSVAPLDRFA-----GSPRYQVAVTEVSGSSLSVSFHLILM 658
DB 543 LSDPANMLKIDPVNGQITTTAVLDRRESIYQNNMNYATFLASDNGIPMSGTGLQIYLL 602
QY 659 DVNDNPPRL-AKDYTGFLFCHPLSAPGSLIFEATDDOHLFRGPHFTFSLSG--SLQND 715
DB 603 DINNAQVNPKEAT---TCETLQ-PNAINITAVDPDIDPNAQ-PAFELPDSPPSIRKN 657

```


TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-332-643-42

Query Match 15.8%; Score 684.5; DB 1; Length 916;
Best Local Similarity 29.3%; Pred. No. 7.3e-52;
Matches 176; Conservative 106; Mismatches 288; Indels 31; Gaps 12;

239 WKAPKEVENENSTDPPIKITQVRW---NDPGAQYSL--VDKEKLPRPPPSID-OEGDI 292
171 WVLP-PINPENSRRGPPQQLVIRSDKNDIPRYSITGVADQPEMEVFSINSMGRM 229
293 YTPQPLDREKDAVYFAVAKDEYKPLSYPLEIHVKVDINDNPPCSPVTFVYQEN 352
230 YVTRPMDREHASYHRAHADVDMNGKNVENPIDLYIVIDMNDHBEFIIQVNCVDEG 289
353 ERGNSIGTLTAHREENTANSFLNRYIVEQPKLPMQGLFIQYAGMLQAKOSLKK 412
290 SKRGTYMTITANDADSTANGVRRIVTQTPQSSQMMFTINSFGDIVVVAAGWDR 349
413 QDPQYVNLIEVSDK-----FKTLCFVOINVINDIQPIPEKSDYGNLTAEPTNIG 466
350 EKVOQYVIVQATDMEGNLNYGLSNTATAIITVDVNDNSEFTASTFAG--EVPENSVE 407
467 STLTITQATDADEPFTGSSKILYHIKGDSEGRGVDTDPHTNGVYIIKKPLDFTAAV 526
408 TVVANLTVMRDQPHSPNMNAVYRIISGDPGHSFVPTDPVTEGMVTVKAVDYELNRA 467
527 SNIVFAKENPEPLVFGKYNAASFAKFTLLIVTDVNEAPOFSOHFAKVEDVAIGTKVG 586
468 FMLTVMVSNQAPLASGIGQMSFOSTAGTISIMDINEAPYPSNHLKRLREGVPPGTVLT 527
587 NVTAKDEGL---DISYSLRGDTRGWLKIDHVTGEIFSVAPLDREA---GSPYRQVVA 639
528 TFSADVDDRFMOQAIVRYSKLSDPASMLHINATNGQITTVAVLDRESLYTKNNVYEATFLA 587
640 TEVGGSSLSVSEPHLLMDVNDNPPRLADYGLFCFCHPLSAPGSLIFEATDDQFLR 699
588 ADNGIIPPASGTGLQIYILIDINDAPLPELPEKAOICERPMLNA---INITAADADVHNT 644
700 GPH-FFTSLSSGSLQNDMEVSKINGTHARLSTHTDEERAYVYLIRINDGRPLEGIV 758
645 GPYVFEPLPVPAAVRKMTITRLNGDYAQLSLILYEAQMYDPIIVTDSGNPLSNTS 704
759 SLPTFSCYEGSCFRPAGHOTGPIYVGMVAGILLTLLVIGIILAV--FIRIKDKXK 816
705 IIKKVCPCDNDGCTTIG---AVAAAGLTGAIVAILICILITLLTWLLFVMMKREK 761
817 D 817
762 E 762

RESULT 8
US-08-332-638-48
Sequence 48, Application US/08332638
Patent No. 5646250
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,638
FILING DATE: 01-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,643
FILING DATE: 17 APR 1992
APPLICATION NUMBER: US/08/049,460
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 5646250and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31340
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 916 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-332-638-48

Query Match 15.8%; Score 684.5; DB 1; Length 916;
Best Local Similarity 29.3%; Pred. No. 7.3e-52;
Matches 176; Conservative 106; Mismatches 288; Indels 31; Gaps 12;

239 WKAPKEVENENSTDPPIKITQVRW---NDPGAQYSL--VDKEKLPRPPPSID-OEGDI 292
171 WVLP-PINPENSRRGPPQQLVIRSDKNDIPRYSITGVADQPEMEVFSINSMGRM 229
293 YTPQPLDREKDAVYFAVAKDEYKPLSYPLEIHVKVDINDNPPCSPVTFVYQEN 352
230 YVTRPMDREHASYHRAHADVDMNGKNVENPIDLYIVIDMNDHBEFIIQVNCVDEG 289
353 ERGNSIGTLTAHREENTANSFLNRYIVEQPKLPMQGLFIQYAGMLQAKOSLKK 412
290 SKRGTYMTITANDADSTANGVRRIVTQTPQSSQMMFTINSFGDIVVVAAGWDR 349
413 QDPQYVNLIEVSDK-----FKTLCFVOINVINDIQPIPEKSDYGNLTAEPTNIG 466
350 EKVOQYVIVQATDMEGNLNYGLSNTATAIITVDVNDNSEFTASTFAG--EVPENSVE 407
467 STLTITQATDADEPFTGSSKILYHIKGDSEGRGVDTDPHTNGVYIIKKPLDFTAAV 526
408 TVVANLTVMRDQPHSPNMNAVYRIISGDPGHSFVPTDPVTEGMVTVKAVDYELNRA 467
527 SNIVFAKENPEPLVFGKYNAASFAKFTLLIVTDVNEAPOFSOHFAKVEDVAIGTKVG 586
468 FMLTVMVSNQAPLASGIGQMSFOSTAGTISIMDINEAPYPSNHLKRLREGVPPGTVLT 527
587 NVTAKDEGL---DISYSLRGDTRGWLKIDHVTGEIFSVAPLDREA---GSPYRQVVA 639
528 TFSADVDDRFMOQAIVRYSKLSDPASMLHINATNGQITTVAVLDRESLYTKNNVYEATFLA 587
640 TEVGGSSLSVSEPHLLMDVNDNPPRLADYGLFCFCHPLSAPGSLIFEATDDQFLR 699
588 ADNGIIPPASGTGLQIYILIDINDAPLPELPEKAOICERPMLNA---INITAADADVHNT 644
700 GPH-FFTSLSSGSLQNDMEVSKINGTHARLSTHTDEERAYVYLIRINDGRPLEGIV 758
645 GPYVFEPLPVPAAVRKMTITRLNGDYAQLSLILYEAQMYDPIIVTDSGNPLSNTS 704
759 SLPTFSCYEGSCFRPAGHOTGPIYVGMVAGILLTLLVIGIILAV--FIRIKDKXK 816
705 IIKKVCPCDNDGCTTIG---AVAAAGLTGAIVAILICILITLLTWLLFVMMKREK 761
817 D 817
762 E 762

```

RESULT 9
US-08-188-228-62
; Sequence 62, Application US/08188228
; Patent No. 5597725
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,228
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/049,460
; FILING DATE: 19 APR 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/872,643
; FILING DATE: 17 APR 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5597725and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31340
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 713 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-188-228-62

Query Match 15.2%; Score 657.5; DB 1; Length 713;
Best Local Similarity 28.7%; Pred. No. 1.2e-49;
Matches 199; Conservative 104; Mismatches 285; Indels 105; Gaps 23;

QY 160 NGQLYQIVQLPMINNVMTFQINNKTKGALSLS--TREGSGELNPAKNPSYN-----LVI 211
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 NDKRLRYEV-----SSPYKXNSDGLVALRNITAVGKTLFVHARTPHADMAELVIV 112
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 212 SVKMGQSEN--SFSSTTSV-----DIYVENIMWAKPKPEVEMENSTDEHPKITTOYRW 264
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 113 GSKDIQSLQDIFFEFARTSPVPROKRSIVS-----PILPENRQPPFRVGCV-- 162
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 265 NDPAQAGSLVDKEKLPRFPF-----SIDOE-----GDIYVQPLDREKDAYVF 308
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 163 -----VSDPEPSKRLTKGVDQEPKGIFRINENTGVSVTRIDREVIANYQL 213
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 309 YAAKDEYKPLSYPLFIHVVKDINDNPTCPSPVTFVEQENBRIGNSIGTLTJARDR 368
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 214 FVEETDVNGKTLLEGVPLEVIVIDQNNRPIFRGPGYIGHWEGSPGTVMRTARDAD 273
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 369 EENFANSFLWRYIVGOTPKLPMDDLFLIQYAG--MLQAKQSLKQOT---PQNLITIEV 424
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 274 DPATDNLRLRNIRKQOTPDKPSPMFYIDPEKGDIVTVSPALLDRETLLENPKYELIEA 333
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

QY 425 SDK-----DEKLCFQVQINVIDINDQIPIFEKSDYGNLTAEPTNIGSTLLTQATDADE 479
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 334 QDMAGLDVGLTGTATATIMIDDKNDSPKFTKKEF--QATVEGAVG--VIVNLTVEDKXD 390
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 480 PFTGSSKILYHIHIGKDESEGLGVDTPHNTGVIYIIKKPLDFTAAVSNIVFAENEPFL 539
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 391 PFTGAMRAAYTIINGNGQGFELIHTNPQTEGMLSVKPLDIYELISAHPTLLIKVENDDL 450
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 540 VFGVKYNASEFAKFTLLVTVDNEAPQPSQVFAKVEDVAIGTKVGNVTAKEPEGD-- 597
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 451 VPDVSYSPSSATATYHTVLDVNEGPPVYPPPMVMTROEDLSVGSVLTVAATPDPSLOH 510
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 598 -ISVSLGDRGMLIKIDHVTGEIJSVAPLDREA---GSPRYQVVAITEVGGSSLSVSSE 652
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 511 TIRSYVKDPAGMINIPINGVDTTAVLDRESPVDNSVYTLAFLADSGNPPATGTGT 570
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 653 PHLLMDVNDNPPRLAKDYGLFPCHPLSAP-----GSLIFEATDDQHLFRGPHFT 704
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 571 LLITLEDVNDNAP-----FIYPTVAEVCDDAKNLSVILIGASDKDLHPNTDP--FK 619
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 705 FSLGGSGSLQND--WEVSKINGTHARLSTRHTDPEERAVVLIIRINDGRPLEGIVSLPYT 763
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 620 FEIHKQAVPKVWKISKINNTHALVSLQ--NLNKNANPLIMVTDGSKPPMTNTDLRVQ 678
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 764 FCSGVBG--SCFRPAGHQTGIPYGVMAVGILLT 794
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 679 VCSGRNSKVDCAAGALRFLSPV-----ILLSS 706
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
US-08-332-643-56
; Sequence 56, Application US/08332643
; Patent No. 5639634
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,643
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/872,643
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5639634and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30795
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 713 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-332-643-56

```


Query Match 15.2%: Score 657.5; DB 1; Length 713;
 Best Local Similarity 28.7%; Pred. No. 1.2e-49;
 Matches 199; Conservative 104; Mismatches 285; Indels 105; Gaps 23;

160 NGOLYYOIVIQLPINNNVWFQINNKTGAISL---TREGSOELNPAKPSYN-----LVI 211
 61 NDKLRIRV-----SSPFKXNSDGLVALNRTNITNGKTLFVHARTPHAEDEMAELIYV 112
 212 SVKMGGOSEN--SFSDTTSV-----DIIVTENIMKAPKEVENENSTDPHPIKITQVRW 264
 113 GKDIOGSLQDIFKFAFRTSPVROKRSIVS-----FILPENORQPFPPDVGVK-- 162
 265 NDPAQVSLVDKELRPFPP-----SIDOE-----GDIVYQPLDREKDAYVF 308
 163 -----VDSRPERSKRFLTGKGVDEPKGIFRINENTGSVSVTRTLDRVIAVYQL 213
 309 YAAKDEYKPLSLPELHIVKVDINDNPTCPSPVTFVEVOENRIGNSIGTLTAHARD 368
 214 FVETTDVNGKTLGEPVLEVIYIDQNDNRIFREGPYIGHMGSPTGTTWMTAFDAD 273
 369 EENTANSFLNRYIVEQTPKLPMDGLFLIQTAYG--MLQAKOSLKKQDT---POYNLTIEV 424
 274 DPATDNLALRYNIRQOTPEKPSPMFYIDPEKGDIVTVVSPALLDRETLLENPKYELIEA 333
 425 SDK-----DKTLCFVOQINVIDINDQPIFEKSDYGNLTLAEDTNGSTLITQATDAE 479
 334 QDMAGLDVGLTGATATATIMDDKNDHSPKTKKEF--QATVEGAVG-VIVNLTVEDKDD 390
 480 PFTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVYIKKPLDPETAASNIVFAKANEPL 539
 391 PTTGAMRAAYTIINGNPGSFEIHTNPOTNEGMLSVKPLDYEISAHTLLIKYENEDPL 450
 540 VFGVKYNAASPFAKFTLITVDYNEAPQESQHVFOAKVSEDVAIGTKVGNVTAKDPEGID-- 597
 451 VPDVSYGSSSTAYVHTVLDVNEGVPYDPDMVTRQEDLSVGSVLLTVNATDSDSIQHQ 510
 598 -ISYSLRGDTGMKIDHVTGEIHSVAPLDREA-----GSPRYQVVAEVEGSSLSVSE 652
 511 TTRSYVKDPAGMININPINGVTDTTAVLDRESPPVDSVYTLAFLAIDSGNPPATGTGT 570
 653 FHILIMDVNDNPPRLADYTGFLFCHPLSAP-----GSLIFEATDDOHLFRGPHFT 704
 571 LLITLEDVNDNAP-----FIYPTVAEVCDDAKNLSVLLGASDKDLHPNTDP--FK 619
 705 FSLGSGSLQND--WEVSKINGTHARLSTRHTDPEERAYVLLIRINDGPRPLEGIVSLPVT 763
 620 FEIHQAVPDYKWKISKINNTHTALVSLIQ--NINKANYMLPIWTDVSGKRPMTNITDLRVQ 678
 764 FCSGVEG--SCFRPAGHOTGIPTYGMAVGILLT 794
 679 VCSGCRNSKYVDCNAGALRFSLSV-----ILLIS 706

RESULT 11

US-08-332-638-62
 Sequence 62, Application US/08332638
 Patent No. 5646250

GENERAL INFORMATION:
 APPLICANT: Suzuki, Shintaro
 TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
 NUMBER OF SEQUENCES: 62
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 STREET: 6300 Sears Tower, 233 S. Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/332,638
 FILING DATE: 01-NOV-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/872,643
 FILING DATE: 17 APR 1992
 APPLICATION NUMBER: US/08/049,460
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5646250and, Greeta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 31340
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 62:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 713 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

Query Match 15.2%: Score 657.5; DB 1; Length 713;
 Best Local Similarity 28.7%; Pred. No. 1.2e-49;
 Matches 199; Conservative 104; Mismatches 285; Indels 105; Gaps 23;

160 NGOLYYOIVIQLPINNNVWFQINNKTGAISL---TREGSOELNPAKPSYN-----LVI 211
 61 NDKLRIRV-----SSPFKXNSDGLVALNRTNITNGKTLFVHARTPHAEDEMAELIYV 112
 212 SVKMGGOSEN--SFSDTTSV-----DIIVTENIMKAPKEVENENSTDPHPIKITQVRW 264
 113 GKDIOGSLQDIFKFAFRTSPVROKRSIVS-----FILPENORQPFPPDVGVK-- 162
 265 NDPAQVSLVDKELRPFPP-----SIDOE-----GDIVYQPLDREKDAYVF 308
 163 -----VDSRPERSKRFLTGKGVDEPKGIFRINENTGSVSVTRTLDRVIAVYQL 213
 309 YAAKDEYKPLSLPELHIVKVDINDNPTCPSPVTFVEVOENRIGNSIGTLTAHARD 368
 214 FVETTDVNGKTLGEPVLEVIYIDQNDNRIFREGPYIGHMGSPTGTTWMTAFDAD 273
 369 EENTANSFLNRYIVEQTPKLPMDGLFLIQTAYG--MLQAKOSLKKQDT---POYNLTIEV 424
 274 DPATDNLALRYNIRQOTPEKPSPMFYIDPEKGDIVTVVSPALLDRETLLENPKYELIEA 333
 425 SDK-----DKTLCFVOQINVIDINDQPIFEKSDYGNLTLAEDTNGSTLITQATDAE 479
 334 QDMAGLDVGLTGATATATIMDDKNDHSPKTKKEF--QATVEGAVG-VIVNLTVEDKDD 390
 480 PFTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVYIKKPLDPETAASNIVFAKANEPL 539
 391 PTTGAMRAAYTIINGNPGSFEIHTNPOTNEGMLSVKPLDYEISAHTLLIKYENEDPL 450
 540 VFGVKYNAASPFAKFTLITVDYNEAPQESQHVFOAKVSEDVAIGTKVGNVTAKDPEGID-- 597
 451 VPDVSYGSSSTAYVHTVLDVNEGVPYDPDMVTRQEDLSVGSVLLTVNATDSDSIQHQ 510
 598 -ISYSLRGDTGMKIDHVTGEIHSVAPLDREA-----GSPRYQVVAEVEGSSLSVSE 652
 511 TTRSYVKDPAGMININPINGVTDTTAVLDRESPPVDSVYTLAFLAIDSGNPPATGTGT 570
 653 FHILIMDVNDNPPRLADYTGFLFCHPLSAP-----GSLIFEATDDOHLFRGPHFT 704
 571 LLITLEDVNDNAP-----FIYPTVAEVCDDAKNLSVLLGASDKDLHPNTDP--FK 619
 705 FSLGSGSLQND--WEVSKINGTHARLSTRHTDPEERAYVLLIRINDGPRPLEGIVSLPVT 763

Db 620 FEIHKQAVPKVWKISKINTHALVSLIQ-NLNKANTNPLMTDSKPEPNTITDLRVQ 678
QY 764 FCSCVEG--SCFRPAGHOTGIPVGMVAILLT 794
Db 679 VCSGRNSKVDGMAAGALRPLSPV-----ILLS 706

RESULT 12

US-08-453-702A-98
Sequence 98, Application US/08453702A
Patent No. 5891706
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 115
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
STREET: 233 South Wacker, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453, 702A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5891706and, Greta B.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32657
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ. ID NO: 98:
SEQUENCE CHARACTERISTICS:
LENGTH: 555 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-453-702A-98

Query Match 15.1%; Score 650.5; DB 2; Length 555;
Best Local Similarity 30.5%; Pred. No. 3,3e-49;
Matches 174; Conservative 87; Mismatches 252; Indels 57; Gaps 13;

QY 239 WKAPKPVENVSTDPHPKITYVRWN-----DPGQYSLVDEKELPRPPEST 286
Db 2 WVIP-PINLPENSRGPPELVRISRDNLRLSVSVTPGA-----DQPPTGIFII 53
QY 287 DQ-EGDIYVTPDLREBKDAYVFAVAVKDEYKPLSYPLEIHKVQKINDNPPTCSPTV 345
Db 54 NPIISGQLSVTKPDLREILAFHRLAHAVDINGNOVENPDIIVINVDMDNDRPEPLHQV 113
QY 346 VFEVQENRIGNSIGTLTAHDREENTANPLNRYIVEQTPKLPMDGLFIQYAGMLQL 405
Db 114 NGSVPESGSKGYVMTVTALIDADDPMALNGMLRYRLISQAPSPSPMFMFINNETGDIIT 173
QY 406 AKQSLKKQDTPQVNLITVEVDK-----PKTLCFQVQINVIDINDQIPF-EKSDYGNLT 458
Db 174 VAAGLREKVOYTLIIQATDMGPNPTYGLSNATRAVITYTVDNDNPEFTAMTFYGEV- 232
QY 459 LAEDTNGISTILITQADDEPFGSSKLIYHIKDGSEGRIGVNDPPHNTGTVIILKPK 518
Db 233 --PENNDVIVAVLTYTDDKQHPITPAMNAIRISGSDPFRKFLIDPNSNDGLVTVVKE 290
QY 519 LDEFYAVSNIVKAEENPELVFGVYKNASSFAKFTLITVDNEADQFSQHVQAKVSED 578

Db 291 IDEFTNMEFVLTVAENQVPLAKGIOHPPOSTATVSVTVADVNEENFYAFNPKITIROEG 350
QY 579 VAIQTVGANTADPEGL---DISYLRGOTRCKMLKDHYTGRTIFSAPLDRRA-----GS 631
Db 351 LHAQTMLTITLADPDRYMOQNIYTKLSDPANMLKIDPVNGQITTTAVLDRESPPVQNN 410
QY 632 PYRQVATEVGGSSLSVSEFHLIMDVNDNPRLAKDYTGLFCHPLSA-----PG 684
Db 411 IVNATPLASNGILPPMSGCTLOIYILDINDNAPQVL-----PQAEFCETPEPN 460
QY 685 SLIFEATDDQHLFRGPH-FTFSLGSGSLQNDWEVSKINGTARLSTRHTDPEERAYVL 743
Db 461 SINIAALDYVIDENAGPFAPDLPLSPVTIKRMTIRLNGDFAQMLKIKFLEAGIYEVP 520
QY 744 IRLNDGRPPLBEGIVSLPTFCSC-VEGSC 772
Db 521 IITDSGNPKPSNISILRVVCCQDSNGDC 550

RESULT 13

US-08-474-067-7
Sequence 7, Application US/08474067
Patent No. 581518
GENERAL INFORMATION:
APPLICANT: Ranscht, Barbara
TITLE OF INVENTION: T-Cadherin Adhesion Molecule
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474, 067
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/213,361
FILING DATE: 14-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/607,293
FILING DATE: 30-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1682
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ. ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 837 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-474-067-7

Query Match 14.7%; Score 634; DB 2; Length 837;
Best Local Similarity 30.4%; Pred. No. 2e-47;
Matches 189; Conservative 104; Mismatches 261; Indels 68; Gaps 25;

QY 239 WKAPKPVENVSTDPHPKITYVRWN---DPGQYSLVDEK-ELKPRPPESTDOE-GDI 292
Db 112 WVIP-PISCLNHRGYPRLVQIKSNKDKSKYYSITGQGADSPVGIIFILERGTWL 170
QY 293 YVQPLDREBKDAYVFAVAVKDEYKPLSYPLEIHKVQKINDNPPTCSPVTFEVQEN 352

```

Db 171 EVTEQDLREKIDRTLLSHAVSAGQGVEDPMEIITVMOQNDKPFIFKEVFGYIEEN 230
QY 353 ERLGNSIGTLTAHDRDEE-NTANSFLNRYIVEQTPKLPMDGLFIQTYAGMLQAKSLK 411
Db 231 AKGTSVMTVATADDAVNTDNGIVSYISQOPRPHQMFIDPAKGIIISVLGTGLD 290
QY 412 KQTPQYNLTIEVSDKDFLT---CFVQINVIDNQIPPEKSDYGNLTAEPTNIGST 468
Db 291 RETTPYVTLVQATDQEGKLSNTATATAIIEVTDANDNIPFNPTMYEG--VVEENKRGTE 348
QY 469 ILTIQATDADEPFTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVIIKKPLDEFTAAVN 528
Db 349 VARLTVDQAPSPAMQAVYHIKSGNDGAFSIITDPSTNGILTKAKGLDYEYTSRYD 408
QY 529 IVFAENPEPLVFGVKNASSFAFCTLIYDVNEAPQFSQVPAQKSEVAIGTKVGNV 588
Db 409 LVVTVENKVPISVPI---TLSTASVLVTVLDVNEPVPFPIKGVGPEDLPGQOVTSY 465
QY 589 TAKDPE---GLDISYSLRGDTRGWLKIDHVTGEIFSVAPLDREA---GSPYRVQVATE 641
Db 466 TAEPDMDROKITYRMGSDPAGWLYHPENGIVTATQPLDRSVHAINSTYKAIITLAVD 525
QY 642 VGGSSLSVSEFHLILMDVNDNPRLAKDYTGTFCHPLSAPGSLIFEATDDQHLERGP 701
Db 526 NGIPDTTGITGLLLLDQVNDNGP--TPERSFEICS--RQPEKQILSYDKD---LPP 577
QY 702 H---FTFSLGSGSLQNDWEYSKINGTHARLSTRHTDPEERAYVVLIRINDGRRPLEG-- 756
Db 578 HTVPFKALAEHGS--SNMWTV--EIRGODELAMGLKKELEPGEVNI FVKLTDS---QGKA 630
QY 757 -IVSLPTFSCVCGE---SCFRPA---GHQGTIPV--GMAVGILTLTLVIGIILAVF 807
Db 631 QVTVAKAQVCEC--EGTAKCERSYIVGG--LGVPAIILIGILGIL---ALLIILLILLF 684
QY 808 IRIKK-----DKGKDNV 819
Db 685 ARRRKVEKEPLLPEDDMRDNV 706

RESULT 14
US-08-474-068A-7
Sequence 7, Application US/08474068A
Patent No. 5837525
GENERAL INFORMATION:
APPLICANT: Ranscht, Barbara
TITLE OF INVENTION: T-Cadherin Adhesion Molecule
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,068A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/213,361
FILING DATE: 14-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/607,293
FILING DATE: 30-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815

```

```

REFERENCE/DOCKET NUMBER: P-LJ 1683
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 837 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-474-068A-7

Query Match 14.7%; Score 634; DB 2; Length 837;
Best Local Similarity 30.4%; Pred. No. 2e-47;
Matches 189; Conservative 104; Mismatches 261; Indels 68; Gaps 25;

QY 239 WKAPKEVWENSTDPPIKITQVKN---DPEAQVSLYDK--EKLPRFPFSIDOE-GDI 292
Db 112 WVIP-DISCLEENRGYPMLVQIKSNKDKESKVYISITGQADSPVGIIFIERETGWL 170
QY 293 YVTPDLREKDAVVEYAAKQEGKPLSYPLEIHVKVDINDNPPTCSPTVTFEVOEN 352
Db 171 EVTEQDLREKIDRTLLSHAVSAGQGVEDPMEIITVMOQNDKPFIFKEVFGYIEEN 230
QY 353 ERLGNSIGTLTAHDRDEE-NTANSFLNRYIVEQTPKLPMDGLFIQTYAGMLQAKSLK 411
Db 231 AKGTSVMTVATADDAVNTDNGIVSYISQOPRPHQMFIDPAKGIIISVLGTGLD 290
QY 412 KQTPQYNLTIEVSDKDFLT---CFVQINVIDNQIPPEKSDYGNLTAEPTNIGST 468
Db 291 RETTPYVTLVQATDQEGKLSNTATATAIIEVTDANDNIPFNPTMYEG--VVEENKRGTE 348
QY 469 ILTIQATDADEPFTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVIIKKPLDEFTAAVN 528
Db 349 VARLTVDQAPSPAMQAVYHIKSGNDGAFSIITDPSTNGILTKAKGLDYEYTSRYD 408
QY 529 IVFAENPEPLVFGVKNASSFAFCTLIYDVNEAPQFSQVPAQKSEVAIGTKVGNV 588
Db 409 LVVTVENKVPISVPI---TLSTASVLVTVLDVNEPVPFPIKGVGPEDLPGQOVTSY 465
QY 589 TAKDPE---GLDISYSLRGDTRGWLKIDHVTGEIFSVAPLDREA---GSPYRVQVATE 641
Db 466 TAEPDMDROKITYRMGSDPAGWLYHPENGIVTATQPLDRSVHAINSTYKAIITLAVD 525
QY 642 VGGSSLSVSEFHLILMDVNDNPRLAKDYTGTFCHPLSAPGSLIFEATDDQHLERGP 701
Db 526 NGIPDTTGITGLLLLDQVNDNGP--TPERSFEICS--RQPEKQILSYDKD---LPP 577
QY 702 H---FTFSLGSGSLQNDWEYSKINGTHARLSTRHTDPEERAYVVLIRINDGRRPLEG-- 756
Db 578 HTVPFKALAEHGS--SNMWTV--EIRGODELAMGLKKELEPGEVNI FVKLTDS---QGKA 630
QY 757 -IVSLPTFSCVCGE---SCFRPA---GHQGTIPV--GMAVGILTLTLVIGIILAVF 807
Db 631 QVTVAKAQVCEC--EGTAKCERSYIVGG--LGVPAIILIGILGIL---ALLIILLILLF 684
QY 808 IRIKK-----DKGKDNV 819
Db 685 ARRRKVEKEPLLPEDDMRDNV 706

RESULT 15
US-08-472-481-6
Sequence 6, Application US/08472481
Patent No. 5863804
GENERAL INFORMATION:
APPLICANT: Ranscht, Barbara
TITLE OF INVENTION: T-Cadherin Adhesion Molecule
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California

```

COUNTRY: United States
 ZIP: 92122
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/472,481
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/213,361
 FILING DATE: 14-MAY-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/607,293
 FILING DATE: 30-OCT-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P-LJ 1686
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 837 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

Query Match	Score	DB 2	Length
14.7%	634	2	837

Matches 189; Conservative 104; Mismatches 261; Indels 68; Gaps 25; Best Local Similarity 30.4%; Freq.NO. 26-4/;

QY 233 WKAPRPVEMVENSTDPHPRIKKTQVRNN--DPCAQVSLVDK--EKLPRFPFESIDQE-CDI 2922

Db 112 WVIP-PISCLHNHRGPRPYMRVLQIKSNKKDKSKVYYSITGGGADSPVGIIFITERETMWL 1700

Qy 293 YVTQPLDREKDAYVFAVAKDEYKPLSTYPLSIHVKKVDINDNPPTCPSPVTVEEIQEN 352

Db 171 EVTEDLDREKIDRYTLLSHAVSASGGPVEDPMEIITVMQDNDKPFPIKEFVGVYTEEN 230

Qy 353 ERLGNSIGTTLTAHDRDEE-NTANSFLNYRIEQTPKLPMDFELIQTVGMLQLAKQSLK 4111

Db 231 AKPGISVMTNATADADAVNTDNGIYSYSISQCPREHPQNTTIDAKGIISVLGTGLD 2900

```
0y      412 KQDPYQNLITIEVSDKDFKT--CFVQINVIDINQIPPEKSDYGNLTIAEDNIGST 468
        ::|||:::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      291 RETPNYTLIVQATDEGKGISNTAFALIEVTDANNIPIFNPTMYEG--VEENKPOTE 348
```

```
OY      469 ILTIQATDADEPFTGSSKILVHIKQDSEGRGVDTDPHTNTGYVIIKKPLDFETAASN 528
       : : | | | | : : : : | | | | : : | | | | : : | | | | : :
DB      349 VARLVTVDOPAGSPMAOAVHIIKSGLDGAFSIITPSTNNGILKTAKGLDYETKSRYP 408
```

```

QY      529  IVFKAENPELVFGVKNASSFAKFTLIVTDVNEAPQSFQHFQAKSEDAVIGTKVNV 588
      :| | | | : : : | | | | : : : | | | : : :
Db      409  LVTVTENKVLSPV---TISTASVAVTVLVDNEPVEVPPIKRVGPEDLVYGQOVTSY 465

```

```
QY      589 TAKDPE---GLDISYSLRGDTKGMWKIDHVTGEIFSVAPLDREA----GSPRYQVVATE 641
        ||::|: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db      466 TAEPDRDMROKITTYRMGSDPAGWLYTHPENGIIVTTOPLBRESVHAINSTKAITLAVD 525
```

QY 642 VGGSSLSVSEFHLIMDVNDNPRLAKDYTGLFCHPLSAGSLFEATDDOHLFRGP 701
| : | | | | | : | | | | |
Db 526 NGIPDTGTGTLILLDQVNDGP--TPRPSRFIS--ROEKQIISVND-----LP 577

```
QY      702 H---FTFSLGSGSLNDWEVSKINGTHARLSTRTHTDPEERAYVLIRINDGRPLEG-- 756
         :| |: | | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db      578 HTYPEKAALEHGS-SNNMTV-ETRGDEIAIAMSILKKRIIEPGEFNITFVKTIDS-----CGKA 630
```

Qx 757 - IVSLPVTFSCVEG---SCRPA-----GHQTGIPV-GMAVGILTLTLVIGIILAVF 807
: : | | | : | : | : | : | : | : | : | : | : |

```

Db      631 QVTQVAQVQCEC-EGTAKCCERRSTVGG--LGPAILGILGIL--ALLTLLLLLLF 684
QY      808 IRIKK-----DKGDNV 819
           | : | | |
Db      685 ARRRKVEKEFLPPEDDARDNV 706

```

Search completed: June 22, 2003, 01:09:33
Job time : 38 secs

THIS PAGE BLANK (uspiu)

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2003, 01:06:35 ; Search time 435 Seconds
(without alignments)
4307.270 Million cell updates/sec

Title: US-10-025-380-1081
Perfect score: 4321
Sequence: 1 MIIQAHLSLCLMLYLATG.....DKGKNVESQAQSEVKEPLRS 832

Scoring table:

BLOSUM62			
Xgapop 10.0	Xgapext 0.5		
Ygapop 10.0	Ygapext 0.5		
Fgapop 6.0	Fgapext 7.0		
Delop 6.0	Delext 7.0		

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xld
-Q=/cgm2_1/USPTO.spool/US10025380/rnatc_17062003_173327_24658/app_query.faeta.1.1031
-DB=N.geneseq_101002 -QFMT=fae+cap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIT=45 -DOCLIT=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USR=US10025380 @CGN 1.1 290 @rnatc_17062003_173327_24658 -NCPU=6 -ICPU=3
-NO MMAP -IARGEDUTRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGCLOG
-DEF_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1:	N.GeneSeq_101002:*
2:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1980.DAT:*
3:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1981.DAT:*
4:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1983.DAT:*
5:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1984.DAT:*
6:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1985.DAT:*
7:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1986.DAT:*
8:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1987.DAT:*
9:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1988.DAT:*
10:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1989.DAT:*
11:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1990.DAT:*
12:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1991.DAT:*
13:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1992.DAT:*
14:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1993.DAT:*
15:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1994.DAT:*
16:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1995.DAT:*
17:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1996.DAT:*
18:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1997.DAT:*
19:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1998.DAT:*
20:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA1999.DAT:*
21:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA2000.DAT:*
22:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA2001A.DAT:*
23:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA2001B.DAT:*
24:	/SID52/gcgdata/geneSeq/geneSeq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4321	100.0	3345	20	AA18166	Human HPT-1 protei
2	4321	100.0	3345	22	AA129511	Human intestinal p
3	4321	100.0	3345	24	ABN67931	Gene #3289 used to
4	4321	100.0	3345	24	AB162444	Colon adenocarcino
5	4321	100.0	3345	24	AB164424	Stomach cancer rel
6	4303	99.6	3654	24	ABN97312	Gene #3810 used to
7	4303	99.6	3654	24	AB162484	Colon adenocarcino
8	4303	99.6	3654	24	AB167667	Oesophagus cancer
9	922.5	21.3	2820	22	AAH98190	Human EST-derived
10	922.5	21.3	2820	22	AAH57507	Human kidney cell
11	908.5	21.0	2848	21	AA337087	Human PRO1340 (UNQ
12	908.5	21.0	2848	22	AA922106	Human PRO1340 cDNA
13	908.5	21.0	2848	22	AA954364	DNA encoding prote
14	908.5	21.0	2848	24	ABK33629	cDNA encoding huma
15	878.5	20.3	2938	23	AA515602	Human cDNA encodin
16	771.5	17.9	3540	22	AAK94573	Human full-length
17	758.5	17.6	516	24	ABK44858	cDNA encoding colo
18	699.5	16.2	2853	24	ABN89391	Human N-cadherin e
19	699.5	16.2	4136	23	AA578503	DNA encoding novel
20	696.5	16.1	2824	24	ABN89350	Chicken N-cadherin
21	696.5	16.1	2824	24	ABN89351	Chicken N-cadherin
22	694.5	16.1	3875	12	AAQ11561	Partial endothelia
23	694.5	16.1	4321	24	AB199593	Mouse ischaemic co
24	684.5	15.8	3048	18	AA785433	Human cadherin-4 c
25	684.5	15.8	3048	18	AA761920	Human cadherin-4 c
26	684.5	15.8	3048	24	ABK83876	Human cDNA differe
27	671.5	15.5	3551	15	AAQ72597	Human HT-1376 cell
28	671.5	15.5	3552	20	AA224623	Human lung tumor a
29	671.5	15.5	3552	21	AA65862	Human lung cancer-
30	671.5	15.5	3552	24	AB149081	Human lung tumour
31	671.5	15.5	3742	22	AA987827	Human late stage o
32	667.5	15.4	3212	24	AB166329	Lung cancer relate
33	667.5	15.4	4125	22	AA58391	Human polynucleoti
34	665.5	15.4	3364	22	AA160177	Human polynucleoti
35	660.5	15.3	2690	18	AA785405	Human cadherin-13
36	660.5	15.3	2690	18	AA761927	Full length human
37	630.5	14.6	2779	13	AAQ24857	Sequence of T-cadh
38	630.5	14.6	3959	13	AAQ24856	Sequence of T-cadh
39	624.5	14.5	3156	23	AA586134	DNA encoding novel
40	611	14.1	14530	24	AA199765	Rat fat 3 protein
41	605	14.0	13960	23	AA587164	DNA encoding novel
42	604	14.0	2768	24	ABN89352	Mouse E-cadherin e
43	604	14.0	2768	24	ABN89353	Mouse E-cadherin e
44	601.5	13.9	12791	23	AB111557	Drosophila melanog
45	599.5	13.9	17282	23	AB111556	Drosophila melanog

ALIGNMENTS

RESULT 1
ID AAX18166 standard; DNA; 3345 BP.

AC AAX18166;
DT 04-MAY-1999 (first entry)

DE Human HPT-1 protein coding sequence.

KX Gastro-intestinal transport receptor; binding protein; hsi; HPT1;

KW DH; hPEPT1; human; GI tract receptor; sucrose-isomaltase complex;

KW intestinal peptide-associated transporter; hypertension; diabetes;

KW osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;

KW therapeutic agent delivery; therapy; probe; ss.

OS Homo sapiens.

XX

QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 DB 1468 GAAGACACAAACATGGGTGCCACCATCTTAACCATCCAGGCCACTGATGCTGATGAACCA 1527
 QY 481 PheThrGlySerSerIleValLeuThrHisIleIleLeuGlyAspSerGluGluValLeu 500
 DB 1528 TTTACTGGAGGTTCTTAATTCGTATCATATCAATAAGGAGACACTGAGGGACGGCTG 1587
 QY 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLeuGlySProLeuAsp 520
 DB 1588 GGGGTTACACAGATCCCATACCAACCCGGAATATCTAATTAATAAGCCCTTGAT 1647
 QY 521 PheGluThrAlaAlaValSerAsnIleValPheLeuAlaGluAsnProGluProLeuVal 540
 DB 1648 TTTGAACACACAGCTGTTCCACATGTGTTCAACACGAAATCTGAGCCCTCTAGT 1707
 QY 541 PheGlyValAlaSerThrAsnAlaSerSerPheAlaIlePheThrLeuIleValThrAspVal 560
 DB 1708 TTTGGTGTGAAGTACAAATGCAAGTCTTTGGCCAAAGTCCACTTATTTGACAGATGTG 1767
 QY 561 AsnGluAlaProGluPheSerGlnHisValPheGlnAlaIleValSerGluAspValAla 580
 DB 1768 AATGAAGCACTCAATTTTCCCAACAGTATTCCAAGCGAAAGTCAAGTGGAGATGACT 1827
 QY 581 IleGlyThrIleValGlyAsnValThrAlaIleAspProGluGluAspIleSerTyr 600
 DB 1828 ATAGGCACTAAAGTGGGCAATGTGACTGCCAAGATCCAAAGGTCCTGCACTAAGTAT 1887
 QY 601 SerLeuArgGlyAspThrArgGlyTyrPLeuIleHisAspHisValThrGlyIleIlePhe 620
 DB 1888 TCACTGAGGGGACACCAAGAGTGGCTTAAATGACCACTGATGCTGATGATCTTT 1947
 QY 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGluValAlaIleThr 640
 DB 1948 AGTGTGGCTTCATTTGACAGAGAACCGGAGTCCATATCGGTTCAAGTGGGCCACA 2007
 QY 641 GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 DB 2008 GAAGTAGGGGGGTCCTTCCTTAAGCTGTGTGAGAGTTCACCTGATCTTAATGATGTG 2067
 QY 661 AsnAspAsnProProArgLeuAlaIleAspTyrThrGlyLeuPhePheCysHisProLeu 680
 DB 2068 AATGACAAACCTCCAGGCTAGCCAGGACTTACACGGGCTGTGCTTCCTCCATCCCTCC 2127
 QY 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheArgGly 700
 DB 2128 AGTGCACCTGGAGTCTCATTTTCGAGGCTAGTATGATATCAACCTTAATTCGGGGT 2187
 QY 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerTyr 720
 DB 2188 CCCCATTTTACATTTTCCCTCGGCAAGTGAAGCTTACAAACACACATCGGAGATTTCCAA 2247
 QY 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaTyr 740
 DB 2248 ATCAATGAGTACTATATCCCACTGTCTACAGGCAACAGACTTGTGAGGAGAGGGGTAT 2307
 QY 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSerLeu 760
 DB 2308 GTGTCTTGATCCGACATCAATGATGGGGGTGGCCACCTTGAAGGCAATGTTCTTTA 2367
 QY 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
 DB 2368 CCAATTACATTCGACGTTGTGTGAGAGAGAGTTGTTCCGGCCAGCAGGTCCACAGACT 2427
 QY 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 DB 2428 GGGATATCCCACTGTGGGCAATGGAGTGGTATCTCTGACCACTTCCTGGGATGGT 2487
 QY 801 IleIleLeuAlaValAlaPheIleArgIleValLeuAspIleGlyIleAspAsnValGlu 820
 DB 2488 ATTAATTTTACAGATGTGTATTCGATTAAGAGAGATTAAGCAAGATTAATGTGAA 2547

QY 821 SerAlaGlnAlaSerGluValIleProLeuArgSer 832
 DB 2548 AGTGTCAAGCATCTGAAGTCAAACTCTGAGAAC 2583

RESULT 2

AA129511 ID AA129511 standard; cDNA; 3345 BP.

AC AA129511;

XX 12-OCT-2001 (first entry)

DE Human intestinal peptide-associated transporter HPT-1 cDNA sequence.

XX Human; immunotherapy; diagnosis; colon cancer; colon tumour;

KW immunogenic; gene therapy; vaccine; colonic cancer; ss.

OS Homo sapiens.

PN WO200149716-A2.

PD 12-JUL-2001.

PF 29-DEC-2000; 2000MO-US35596.

XX 30-DEC-1999; 99US-0476296.

PR 10-JAN-2000; 2000US-0480321.

PR 15-FEB-2000; 2000US-0504629.

PR 06-MAR-2000; 2000US-0519444.

PR 19-MAY-2000; 2000US-0575251.

PR 29-JUN-2000; 2000US-0609448.

PR 28-AUG-2000; 2000US-0649811.

XX (COR1-) CORIXA CORP.

PA Xu J, Lodes MJ, Secret H, Benson DR, Meagher MJ, Stolk JA;

PI King GE, Wang T, Jiang Y;

XX WPI; 2001-441847/47.

DR Colon tumor associated proteins and nucleic acids useful for the

XX PT prevention, diagnosis and treatment of colonic cancer -

XX PS Claim 2; Page 466-467; 472pp; English.

XX CC The present invention describes colon tumour associated proteins (I) and

CC CC the polynucleotides (II) that encode them. (I) have cytostatic activity.

CC CC (I) and (II) can be used in gene therapy and vaccine production. (I) and

CC CC (II) may be used in the prevention, diagnosis and treatment of diseases

CC CC associated with inappropriate colon tumour associated protein (TCAP)

CC CC expression, such as colonic cancer. For example, (I) and (II) may be

CC CC used to treat disorders associated with decreased expression by

CC CC rectifying mutations or deletions in a patient's genome that affect the

CC CC activity of TCAPs by expressing inactive proteins or to supplement the

CC CC patients own production of them. Additionally, (II) may be used to

CC CC produce the TCAP proteins, by inserting the nucleic acids into a host

CC CC cell culturing the cell to express the protein. (II) and its

CC CC complementary sequences may also be used as DNA probes in diagnostic

CC CC polymerase chain reaction (PCR) and hybridisation assays to detect and

CC CC quantitate the presence of similar nucleic acids in samples, and

CC CC therefore which patients may be in need of restorative therapy. (I) may

CC CC also be used as antigens in the production of antibodies against TCAPs

CC CC and in assays to identify modulators of TCAP expression and activity.

CC CC Anti-(I) antibodies and antagonists may also be used to down regulate

CC CC TCAP expression and activity. The anti-(I) antibodies may also be used

CC CC as diagnostic agents for detecting the presence of TCAPs in samples

CC CC (e.g. by enzyme linked immunosorbent assay (ELISA)). AA128460 to AA129512

CC CC and AA124494 to AA124523 represent nucleotide and amino acid sequences

CC CC given in the exemplification of the present invention.

XX SQ Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;

Alignment Scores:

Prod. No.: 0 Length: 3345
 Score: 4321.00 Matches: 832
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-10-025-380-1081 (1-832) x AA129511 (1-3345)

QY 1 MetIleuGlnAlaHisLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
 DB ATATATCTTCAGGCCCATCTTCACTCCCTGTCCTTCTTATGCTTATTTGGCACTGGA 147
 QY 21 TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGlu 40
 DB TATGGCCAAAGAGGGAAAGTTAGTGAACCCCTGAACCCATGACATTTCTATTATGA 207
 QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheValAlaSerProAlaValThrPhe 60
 DB GGCCAAAGACCGAGTCAATTAATTCAGTTTAAGCCAACTCTCTGCTGACCTTTT 267
 QY 61 GluLeuThrGlyGlnThrAspAsnIlePheValIleGluLysGluGlyLeuLeuTyrTyr 80
 DB GAATTAATCTGGGAGAGACAGACATATTTGTATGAACGGAGGAGACTTGTATTAAC 327
 QY 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 DB AACAGACCTTGGACAGGAAACAAAGATCTACTCAATCTCCAGTTTGGACCCCTGGAC 387
 QY 101 AlaAsnGlyIleIleValIleGluGlyProValProIleThrIleGluValLysAspIleAsn 120
 DB GCTATAGAAATTAATAGAGAGGTCAGTCCCTTACCATAGAAAGTGAAGACATCAAC 447
 QY 121 AspAsnArgProThrPheLeuGlnSerLysTyrGlyLysValAlaGlnAsnSerArg 140
 DB GACATAGACCCCATCTTCTCAGTCAAGTACAGAGGCTCAGTAAGGACAACTCTGCG 507
 QY 141 ProGlyLysProPheLeuTyrValAlaAsnAlaThrAspLeuAspProAlaThrProAsn 160
 DB CCAGAAAGCCCTTCTGTATGTCATGACCAAGCTCGATGATCCGGCCACCTCCCAAT 567
 QY 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
 DB GGCCAGCTTATATCCAGATTGTCATCCAGCTTCCATGATCAACATGTCATGATCTT 627
 QY 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
 DB CAGATCAACAACAAAGGAGGACATCTCTTACCCGAGAGGATCTCAGGAATTGAT 687
 QY 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlyGlnSer 220
 DB CCTGCTAAGAAATCTTCTATATATCTGTGATCTCAGTGAAGGACATGGGAGGCAAGT 747
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValIleThrGluAsnIleTyrLys 240
 DB GAGAAATCTTCAAGTATCAACATCTGTGATATCATAGTACAGAGAAATTTGGAAA 807
 QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisproIleLysIleThr 260
 DB GCACCAAAACCTGTGAGATGTGGAAACCTCACTGATCTCAACCCCATCAAAATCACT 867
 QY 261 GlnValAlaGlyTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGluLysLeuPro 280
 DB CAGGTGGGTGGTGAATGATCCGGTGCACAAATATTCCTTAAGTTGACAAAGAAACCTGCA 927
 QY 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrValIleThrGlnProLeuAspArg 300
 DB AGATTTCCATTTCAATTGACACGAGAGGAATATTTTACGTGACTCAGCCCTTGGACCA 987
 QY 301 GluGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu 320
 DB GAAGAAAGAGATGATATGTTTATATGCAAGTGAAGAGATGAGTACGAAACCACTT 1047

QY 321 SerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProThrCys 340
 DB TCATATCCCTGGAAATTAATGTAAGTAAAGATTAATTAATGATATACCACTACATGT 1107
 QY 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 DB CCGTCAACAGTAACCGTATTTGAGTCCAGAGAAATGAACGACTGGGTAAACGATACCGG 1167
 QY 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
 DB ACCCTTACGTGACATGACAGAGGATGAAGAAATACCTGCCAACAGTTTCTTAACCTCAGG 1227
 QY 381 IleValIleGlnIleThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 DB ATTTGTGAGCAACTCCCAACTCCCATGAGATGACTCTTCTTATCCAAACCTATGCT 1287
 QY 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysValGlnAspThrProGlnTyrAsnLeu 420
 DB GGAAATGTACAGTTACTTAACAGTCTTGAAGAAACAAAGATACTCTCAGTACAACTTA 1347
 QY 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
 DB ACGATAGAGGTGTCTGACAAAGATTTCAAGACCTTTGTGTGTGCAAAATCAACGTTAT 1407
 QY 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAla 460
 DB GATATCAATGATAGATCCCATCTTGTGAATAATCAATATGAAACCTGACTTGTCT 1467
 QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 DB GAAGACAAACAACTTGGGTCCACATCTTAACCATCAGGCCACTATGCTGATGAGCCA 1527
 QY 481 PheThrGlySerSerLysIleLeuTyrHisIleIleValGlyLysAspSerGluTyrArgLeu 500
 DB TTTTACTGGAGTCTAAATCTGTATCATATCAATTAAGGAGAACAGTGGAGCGCTG 1587
 QY 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
 DB GGAGTTGACACAGATCCCATCCATCCAAACCGGATATGTCATATTAAGCCCTTGAT 1647
 QY 521 PheGluThrAlaAlaValSerAsnIleValPheValAlaGluAsnProGluProLeuVal 540
 DB TTTGAAACAGCAGCTGTTTCCATCAATGTGTGTCAAAGACGAAATCTGAGCTCTAGTG 1707
 QY 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValIleThrAspVal 560
 DB TTTGGTGTGAAGTACAAATGCAAGTCTTTTGGCAAGTTCAGCTTATTTGACAGATGTG 1767
 QY 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 DB AATGAGCACTCAATTTTCCCAACAGTATTCCAAGCGAAATCAAGTACAGTACATAGCT 1827
 QY 581 IleGlyThrLysValGlyAsnValIleThrAlaLysAspProGluLysAspIleSerTyr 600
 DB ATAGGCACTAAAGTGGCAATGTGATGCTGCACAAAGATTCAGAAAGTGCAGATAGCTAT 1887
 QY 601 SerLeuArgGlyAspThrArgGlyTyrPheLysIleAspHisValIleThrGlyIlePhe 620
 DB TCACTGAGGGGAAACCAAGAGGTGGCTTAAATGATGACCAAGTACGATGGTGGATCTT 1947
 QY 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaAlaThr 640
 DB AGTGTGCTCATTTGACAGAGAGCCGGAAGTCCATATGGGTACAAAGTGGTGGCCACA 2007
 QY 641 GluValGlyLysSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 DB GAAGTAGGGGGCTTCTTAACTGCTGATCAGAGTCCACATGATCTTATAGAGATGTG 2067
 QY 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 DB AATGACAAACCTCCAGAGCTAGCCAAAGACTTACAGGGCTTGTCTTCTCCATCCCTTC 2127
 QY 681 SerAlaProGlySerLeuIlePheGluAlaIleThrAspAspAspGlnHisLeuPheArgGly 700

Db 2128 AGTCACCTGGAAATCCATTTCGAGGCACTATGATGATGACACTTAATTCGGGGT 2187
 Qy ProtiapherThrPheSerLeuGlySerGlySerLeuGlnAsnAPTPGLuValSerIlys 720
 Db 2188 CCCATTTCATTTCCTTCCTGGGAGTGGAAAGCTTCAAAACGACTGGAGATTCCAAA 2247
 Qy 721 ILeAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluArgAlaTyr 740
 Db 2248 ATCAATGGTACTACATGCCGACCTGTCTACCAAGCACACAGCTTTGAGAGAGGGCGAT 2307
 Qy 741 ValValLeuIleAlaGlyIleAsnAspGlyValArgProProLeuGluGlyIleValSerLeu 760
 Db 2308 GTGCTCTTGATCCCGCATCAAGATGGGGGCTGGCACCCCTTGAAAGCATTTGTTCTTTA 2367
 Qy 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
 Db 2368 CCAATTACATTCCTCAAGTTGTGTGGAAAGAGATTGTTCCGGCCAGAGGTCACCAAGCT 2427
 Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2428 GGGATACCACTGTGGGCGATGGCAGTTGGTATACCTGCGACCACTTCCTGGTATGGT 2487
 Qy 801 IleIleLeuAlaValValPheIleArgIleIleValAspIysGlyIleAspAsnValGlu 820
 Db 2488 ATATATTTACAGATGTGTGTATTCGCAATAAAGAGATAAAGCAAGATTAATGTGAA 2547
 Qy 821 SerAlaGlnAlaSerGluValIlePheProLeuArgSer 832
 Db 2548 AGTCTCAAGCATCTGAAGTCAACCTCTGAGAAAGC 2583
 RESULT 3
 ABN96791 ID ABN96791 standard; DNA; 3345 BP.
 XX
 AC ABN96791;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Gene #3289 used to diagnose liver cancer.
 XX
 KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 KW metastatic liver tumour; cytostatic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 OS Homo sapiens.
 XX
 PN WO200229103-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 02-OCT-2001; 2001WO-US30589.
 XX
 PR 02-OCT-2000; 2000US-237054P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 XX
 XI WPI; 2002-426119/45.
 DR
 XX
 PT Diagnosing and detecting the progression of liver cancer,
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample
 XX
 PS Claim 1; SEQ ID NO 3289; 298bp; English.
 XX
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumor in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a

CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WPIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;
 Alignment Scores:
 Pred. No.: 0 Length: 3345
 Score: 4321.00 Matches: 832
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 US-10-025-380-1081 (1-832) x ABN96791 (1-3345)
 Qy 1 MetIleLeuGlnAlaHisIleuHisSerLeuCysLeuLeuMetIleuValIleThrGly 20
 Db 88 ATGATATCTCAGGCCCATCTTCACTCCCTGTGCTTCTTATGCTTATTTGGCAACTGGA 147
 Qy 21 TyrGlyGluGluGluIlePheSerGlyProLeuIleProMetThrPheSerIleTyrGlu 40
 Db 148 TATGGCCCAAGAGGGAGATTATGTGAGCCCTTAAACCCCAATGACATTTCTATTATGAAG 207
 Qy 41 GlyGluGluProSerGlnIleIlePheGlnPheAlaAsnProProAlaValIleThrPhe 60
 Db 208 GGCAGAAACCGAGTCAATATATTCAGTTTAAGCCCATCTCCTGCTGACCTTT 267
 Qy 61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluArgGlyGlyLeuLeuTyrTyr 80
 Db 268 GAACATACTGGGAGACAGACACATATTTGTGATGAAGACGGAGGACTTGTATTAAC 327
 Qy 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuIleValIleAlaIleAsp 100
 Db 328 AACAGAGCCTTGACAGGAGAAACAAATCTACACAAATCTCAGGTTGACGCCCTGGAC 387
 Qy 101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValIleAspIleAsn 120
 Db 388 GCTAATGGAAATTAATGTGAGGGTCCAGTCCCTATACCACTAAGATGAAGACATTAAC 447
 Qy 121 AspAsnArgProThrPheLeuGlnSerIleTyrGlyGlySerValIleArgIleAsnSerArg 140
 Db 448 GACAAATGCACCCAGTTTCTCCAGTCMAAGTACGAAGGCTCACTAAGGACGAACCTCGC 507
 Qy 141 ProGlyIleProPheLeuTyrValAlaAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
 Db 508 CCAGAAAGGCCCTTCTGTATGCAATGCCAGACACTGTGATATCCGGCACTCCCAAT 567
 Qy 161 GlyGluLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnValMetTyrPhe 180
 Db 568 GGCACGCTTATTAACAGATGTGCATCCAGCTTCCATGATCAACATGTCACTACTTT 627
 Qy 181 GlnIleAsnAsnIleThrGlyAlaIleSerLeuThrArgGluGlySerGlnIleAsn 200
 Db 628 CAGATCAACAAACAAAGGAGCAGCTCTCTTAACCGAGAGGATCTCAGAAATGAAT 687
 Qy 201 ProAlaIleAsnProSerTyrAsnLeuValIleSerValIleAspMetGlyGlnSer 220
 Db 688 CTGTCTAAGATCTCTTAATATCTGTGATCTCGTAAGACATGGAGGCCAGAGT 747
 Qy 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrPlys 240
 Db 748 GAGAAATTCCTTCAGTATACCATCTGTGATATATGTGACAGAAATATTTGGAA 807
 Qy 241 AlaProIleProValGluMetValGluAsnSerThrAspProIleIleTyr 260

Db 808 GCACCAAAACCTGTGAGATGTGGAAAACTCAACTGATCTCAACCCCATCAAAATCACT 867
 Qy 261 GlnValArgThrPheAspProGluValAlaGlnTyrSerLeuValAspGluLeuPro 280
 Db 868 CAGGTGGCGTGAATGATCCGGTGCACAAAATTCCTTAAGTTGCAAAAGAAACCTGCAC 927
 Qy 281 ArgPheProPheSerIleAspGlnGluValAspIleTyrValThrGlnProLeuAspArg 300
 Db 928 AGATTCCTCATTTCAATTGACCAAGAGATATTTCGTGACTCAAGCCCTTGACCGCA 987
 Qy 301 GluGluIysAspAlaTyrValPheTyrAlaValAlaIysAspGluTyrGlyLeuProLeu 320
 Db 988 GAAGAAAGATGATATGATTTTATATGCAAGTGAAGATGATGCGGAAACCACTT 1047
 Qy 321 SerTyrProLeuGluIleHisValIysValIysAspIleAsnAspAspProProThrCys 340
 Db 1048 TCATATCCGTGGAAATTCATGTAAGTTAAAGTATTAATGATTAATCCCACTCAACGT 1107
 Qy 341 ProSerProValThrValPheGluValAlaGlnIysAspGluTyrGlyLeuValSerIleGly 360
 Db 1108 CCGTACCAAGTAAACCGTATTGAGGTCCAGAGAAATGACACATGGGTAAACGATGCGG 1167
 Qy 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
 Db 1168 ACCCTTACTGCACATGACAGGGATGAAGAAATACCTGCCAACAGTTTCTTAACCTACAGG 1227
 Qy 381 IleValGlnGlnThrProIysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 Db 1228 ATTGTGGAGCAAACTCCCAAACTTCCCATGATGACCTTCTCAATCCAAACCTAATGCT 1287
 Qy 401 GlyMetLeuGlnLeuAlaIysGlnSerLeuIysGlnAspThrProGlnTyrAsnLeu 420
 Db 1288 GGAATTTTACGATTAGCTTAACAGCTCTTGAAAGACAAAGATCTCTCCATCACTACACTTA 1347
 Qy 421 ThrIleGluValSerAspIysAspPheIysThrLeuCysPheValGlnIleAsnValIle 440
 Db 1348 ACGATAGAGGTGTCTGACAAAGATTTCAGAACCTTTGTTTGCAAAATCAAGTATTT 1407
 Qy 441 AspIleAsnAspGlnIleProIlePheGluIysSerAspTyrGlyAsnLeuThrLeuAla 460
 Db 1408 GATATCAATGATCAGATCCCATCTTTGAAAATCAAGATTAAGAAACCTAATCACTTGTCT 1467
 Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db 1468 GAAGACCAAACTTTGGGTCCACCATCTTAACATCCAGGCCACGTGAGCTGAGAGGCA 1527
 Qy 481 PheThrGlySerSerIysIleLeuTyrHisIleIleIysGlyAspSerGluGlyArgLeu 500
 Db 1528 TTTACTGGAGTCTTAATAATCTGATCATATCAATAAGGAGACAGTGAAGGAGCGCTG 1587
 Qy 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleIysLeuProLeuAsp 520
 Db 1588 GGGGTGACACAGATCCCATTAACCAACCGAGATATGTCATTAATTAATAAACCTCTTGAT 1647
 Qy 521 PheGluThrAlaAlaValSerAsnIleValPheIysAlaGluAspProGluProLeuVal 540
 Db 1648 TTTGAAACAGACAGCTTTTCCACATCTGTCTCAAGAGAAATCTCGAACCTCTAGTG 1707
 Qy 541 PheGlyValIysTyrAsnAlaSerSerPheAlaIysPheThrLeuIleValThrAspVal 560
 Db 1708 TTTGGTGTGAAGTACATGCAAGTCTTTTGCAAGTTCACGCTTATTTGACAGATGTG 1767
 Qy 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaIysValSerGluAspValAla 580
 Db 1768 AATGAAGACCTCAATTTTCCCAACACATATTCCAAAGGAAAGTCAGAGAGATGTAGCT 1827
 Qy 581 IleGlyThrIysValIysAsnValThrAlaIysAspProGluGlyLeuAspIleSerTyr 600
 Db 1828 ATAGGACCTTAAGTGGCAATGTGACTGCAAGGATCCAGAAAGTCTTGACATTAAGCTAT 1887
 Qy 601 SerLeuArgGlyAspThrArgGlyTyrPheIysIleAspHisValThrGlyGluIlePhe 620
 Db 1888 TCACGTGAAGGAGACAAAGAGTGTGGCTTAATAATTGACCAAGTGTGAGATCTTTT 1947

Qy 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaThr 640
 Db 1948 AGTGTGCTTCATTTGGACAGAGAAACCGAGATCCATATCGGTACAAAGTGGCCACA 2007
 Qy 641 GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 Db 2008 GAAGTAGGGGGGTCTTCTTAAGCTCTGTGTCAAGATTCACCTGATCTTATGAGATGTG 2067
 Qy 661 AsnAspAsnProProArgLeuAlaIysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 2068 AATGACAAACCTCCAGGCTAGGCAAGACCTACACGAGCTTGTTGTTCTGCAATCCCTC 2127
 Qy 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspHisIleuPheArgGly 700
 Db 2128 AGTGCACCTGGAAGTCTCATTTTTCAGAGCTACTGATGATATACACATTAATTCGGGT 2187
 Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerIys 720
 Db 2188 CCCCATTTTACATTTTCCCTCGGACATGGAAAGCTTCAAAAAGACTGGGAAAGTTCCAAA 2247
 Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaTyr 740
 Db 2248 ATCAATGTGACTGATGCGGACACTGTCTACACAGGACACACACTTTGAGAGAGCGGTAT 2307
 Qy 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSerLeu 760
 Db 2308 GTGCTCTTGAATCCGCATCAATGATGGGGGTGGCCACCTTGAAAGCATGTTTCTTTA 2367
 Qy 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlnHisGlnThr 780
 Db 2368 CCAAGTACATTTGCAAGTGTGTGGAAGAAAGTTGTTCCGGCAGAGGCTCACCAAGACT 2427
 Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2428 GGGATACCCCACTGTGGGCATGAGCAAGTGTGATCTGTCACACCTTGTGTGATGTGT 2487
 Qy 801 IleIleLeuAlaValAlaPheIleArgIleIysLeuAspIysGlyIysAspAsnValGlu 820
 Db 2488 ATATTTTACAGATGTTGTTTATCCGCTAAAGAGATTAAGCAAAAGTATATGTTGAA 2547
 Qy 821 SerAlaGlnAlaSerGluValIysProLeuArgSer 832
 Db 2548 AGTGTCAACATCTGAAGTCAAACTCTGAGAAGC 2583
 Db 2583

RESULT 4
 ABL62444
 ID ABL62444 standard; DNA; 3345 BP.
 AC ABL62444;
 XX
 DT 15-MAY-2002 (first entry)
 DE
 XX
 DE Colon adenocarcinoma related gene sequence SEQ ID NO: 781.
 DE
 XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 OS
 XX
 PN W02001.94629-A2.
 PN
 XX
 PD 13-DEC-2001.
 PD
 XX
 PF 30-MAY-2001; 2001WO-US10838.
 PF
 XX 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.

PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 (AVAL-) AVALON PHARM.

PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,
 PI Soppet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.

PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set

XX Claim 1; SEQ ID 781; 44pp; English.

XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (1) of a signature gene set, where (1)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (1) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.

XX Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;

Alignment Scores:

Pred. No.: 0 Length: 3345

Score: 4321.00 Matches: 832
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

US-10-025-380-1081 (1-832) x ABL62444 (1-3345)

QY 1 MetIleuGlnIahIstleuHISserIeuCYaleuLeuMetIeuTYrleuaIatHrGly 20
 DB 88 ATGATCTTCAAGCCCATCTTCACTCCCTGTGTCTTATGCTTATTTGGCAACTGGA 147
 QY 21 TYrGIyGlnIuGIyusPheSerGIyProIeuLYsProMetThrPheSerIleTYrGIu 40
 DB 148 TATGGCAAGAGGGGAAGTTTATGTGACCCCTGAACCATGACATTTCTATTATGA 207
 QY 41 GYlGhIuPProSerGlnIleIlePheGlnPheLYaIAsnDProAlaValThrPhe 60
 DB 208 GGCCAAAGAACCGAGTCAAAATTATATTCAGTTTAAAGGCCAATCTCTGCTGACTTTT 267
 QY 61 GlnIeuThrGIyGIuThrAspAsnIlePheValIleGIuAArgGIyIleuIeuTYrTYr 80
 DB 268 GAACCTAACCTGGGAGACAGACACATATTTGTATGAAAGGAGGACTTGTATATAC 327
 QY 81 AsnArgAlaIeuAspArgIuThrArgSerThrHisAsnIeuGlnValAlaAlaIeuAsp 100
 DB 328 AACAGAGCCTTGGACAGGGAAACAAATCTACTCAAACTCCAGTTGAGCCCTGGAC 387
 QY 101 AlaAsnGIyIleIleValGIuGIyProValProIleThrIleGIuValIyAspIleAsn 120
 DB 388 GCTAATGGAATTAATGAGGGGTCCAGTCCCTATACCATAGAGTAAGAGACATCAAC 447
 QY 121 AspAsnArgProThrPheIeuGlnIeuIleValIleGlnIeuProMetIleAsnValMetTYrPhe 140
 DB 448 GACAAATGACCCACGTTTCTCCAGTCAAGTAAGAGGCTCAATAGGACGAATCTCCG 507
 QY 141 ProGIyLYsProPheIeuTYrValAsnAlaThrAspIeuAspProAlaThrProAsn 160
 DB 508 CCAAGAAAGCCCTCTTCTGTATGTCATGCAACGACCTGTATATCCGGCATCTCCAA 567
 QY 161 GYlGlnIeuTYrTYrGlnIleValIleGlnIeuProMetIleAsnValMetTYrPhe 180
 DB 568 GGCCAGCTTATATACAGATTCATGATTCATCCAGCTTCCCATGATCAACATGTCATG 627
 QY 181 GlnIleAsnAsnIeuThrGIyAlaIleSerIeuThrArgGIySerGlnIeuAsn 200
 DB 628 CAGATCAACAAACAAACGGAGGCATCTCTTAAACCGAGAGGATCTCAGAAATGAA 687
 QY 201 ProAlaLYsAsnProSerTYrAsnIeuValIleSerValIyAspMetGIyGIuIns 220
 DB 688 CTGTCTAAGATCTCTTCTTATATCTGTATCTCGTGAAGACATGGAGGCGCAAGT 747
 QY 221 GlnAsnSerPheSerIeuThrSerValAspIleIleValThrGIuAsnIleTYrLYs 240
 DB 748 GAGAAATCTCTTACGATATCCACATCTGTGATATCATGTGACAGAAATATTGAAA 807
 QY 241 AlaProLYsProValGIuMetValGIuAsnSerThrAspProHisProIleLYsIleTr 260
 DB 808 GAACCAAAACCTGTGAGATGTGGAAACCTCAACTGATCCCATCAAAATCACT 867
 QY 261 GlnValArgTrpAsnAspProGIyAlaGlnTYrSerIeuValAspLYsGIuLYsLeuPro 280
 DB 868 CAGGTCCGTGATATATCCGGTGCACATATCTTTAGTTGACAAAGAGAGCTGCCA 927
 QY 281 ArgPheProPheSerIleAspGlnIuGIyAspIleTYrValThrGlnProIeuAspArg 300
 DB 928 AGATTCCATTTTCAATATGACAGAGAGATATTATGATGATCAGCCCTTGACCGA 987
 QY 301 GlnGIuLYsAspAlaTYrValPheTYrAlaValAlaLYsAspGIuTYrGIyLYsProLeu 320
 DB 988 GAAGAAAGATGATATGTGTATTTATATGATGCAAGAGATGAGTAAGCAAAACCACTT 1047
 QY 321 SerTYrProIeuGlnIleHisValIyAspIleAsnAspAsnProProThrCYs 340

Db 1048 TCATATCCCGGAAATTCATGTAATAAGTTAAAGTATTAATGATATTCACCTACACTGT 1107
 Qy 341 ProserProvalThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 Db 1108 CCGTCAACAGTAAACCGTATTGAGGTCCAGAGAAATACACACGCGGTAAACAGTATCGGG 1167
 Qy 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnThrArg 380
 Db 1168 ACCCTTACTGCACATGACAGGAGTGAAGAAATACTGCACACAGTCTTCTAACTACAGG 1227
 Qy 381 IleValGluGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 Db 1228 ATTGTGAGCAAACTCCCAAACTTCCCATGATGAGCTCTTCCATTCAACTAAGCT 1287
 Qy 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAsnThrProGlnThrAsnLeu 420
 Db 1288 GGAATGTTACAGTAACTAAACAGTCCCTGAAAGACAGATACCTCTCAGTACAACTTA 1347
 Qy 421 ThrIleGluValSerAspLysAspPheLysThrLeuGlyAsnValIle 440
 Db 1348 ACGATTAAGGTGTCTGACAAAGATTTCAGACCTTGTGTTTGTGCAAAATCAAGTAACT 1407
 Qy 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAla 460
 Db 1408 GATATCAATGATCAGATCCCATCTTTGAAATAATCAGATTATGAAACCTGACTTGTCT 1467
 Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db 1468 GAAACACAAACATTTGGGTCCACCATCTTAACTCCAGGACCTGATGCTGATGAGACCA 1527
 Qy 481 PheThrGlySerSerLysIleLeuThrHisIleIleLysGlyAsnSerGluArgLeu 500
 Db 1528 TTATCTGGAGATCTTAATAATCTGATCATATCAATTAAGGAGACAGAGAGGAGGCTG 1587
 Qy 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
 Db 1588 GGGGTGACACAGATCCCATACCAACACCGGATATGTCATTAATAAAGCCTTGAT 1647
 Qy 521 PheGluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 Db 1648 TTTGAAACAGACAGCTGTTCCAACTGTTGTTCAAGACAGAAATCCCTGAGCCTCTAGTG 1707
 Qy 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 Db 1708 TTTGCTGGAAGTCAATGCAAGTCTTTTGCAGATTCACGCTTATTTGACAGATGTG 1767
 Qy 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 Db 1768 AATGAGACCTCAATTTTCCCAACACGATATTCAGAGGAAAGTCAGTGAAGATGAGCT 1827
 Qy 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluLysLeuAspIleSerTyr 600
 Db 1828 ATAGGCACTTAAGGCGCATGTGACTGCCAAGGATCCAGAGGCTGCGACATAGCTAT 1887
 Qy 601 SerLeuArgGlyAspThrArgGlyThrLeuLysIleAspHisValThrGluIlePhe 620
 Db 1888 TCACCTGAGGAGACACAGAGGTTGCTTAAATTTGACACGAGTCTGGGAGATCTTT 1947
 Qy 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaThr 640
 Db 1948 AGTGTGCTCATTTGACAGAGAGCCGGAAGTCCATTCGGGTACAGAGTGTGGCCACA 2007
 Qy 641 GluValGlyLysSerSerLeuSerValSerGluPheHisLeuIleLeuMetAspVal 660
 Db 2008 GAAATGAGGGGGGTCTTCTTAAGCTCTGTGTCAAGTTCACACCTGATCTTATGAGTGTG 2067
 Qy 661 AsnSerAsnProArgGluValAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 2068 AATGACAACTCCCAAGGCTGACAGAGACTACAGGCGGTGTTCTTTCGCAATCCCTC 2127
 Qy 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspArgGlnHisLeuPheArgGly 700

Db 2128 AGTGACCTGGAAGTCTCATTTTTCAGGCTACTGATGATGATCAGCACTTAATTCGGGGT 2187
 Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGluAsnAspThrGluValSerLys 720
 Db 2188 CCCCATTATTAACATTTTCCCTCGGACAGTGAAGCTTACAAACACAGTGGAAAGTTCCAAA 2247
 Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaTyr 740
 Db 2248 ATCATGTGATCATATCCCGACATGCTTACCAAGCACACAGACTTTAGGAGAGGGGTAT 2307
 Qy 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGlnGlyIleValSerLeu 760
 Db 2308 GTGCTTGTATCCGCAATCATATGAGGGGCGGCCACCTTGGAAAGCAATGTTTCTTA 2367
 Qy 761 ProValThrPheCysSerCysValGlyLysSerCysPheArgProAlaGlyHisGlnThr 780
 Db 2368 CCAATTACATCTGCAAGTGTGTGGAAGAAAGTGTTCGGGACACAGGTCCACAGACT 2427
 Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuValIleGly 800
 Db 2428 GGGATACCACTGTGGGCAATGGCAGTTGGATATCTGTCACACCTTCTGTGATTTGGT 2487
 Qy 801 IleIleLeuAlaValAlaPheIleArgIleLysLysAspLysGlyLysAsnValGlu 820
 Db 2488 ATAAATTTAGCAGTTGTGTTATCCCATTAAGAAAGATTAAGGCNAAGTATATGTTGAA 2547
 Qy 821 SerAlaGluAlaSerGluValLysProLeuArgSer 832
 Db 2548 AGTGCTCAAGCATCTTAAGTCAAACCTCTGAGAACG 2583
 RESULT 5
 ID ABL64424 standard; DNA; 3345 BP.
 XX ABL64424;
 AC
 DT 15-MAY-2002 (first entry)
 XX
 DE Stomach cancer related gene sequence SEQ ID NO:2761.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytosarctic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001MO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 22-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 26-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.

PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 XX (AVAL-) AVALON PHARM.
 XX
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 XX WPI; 2002-188264/24.
 XX
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 XX Claim 1; SEQ ID 2761; 44pp; English.

XX
 XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.

XX
 XX Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;

Alignment Scores:

Pred. No.: 0 Length: 3345
 Score: 4321.00 Matches: 832
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

US-10-025-380-1081 (1-832) x ABL64424 (1-3345)

QY 1 MetTLeuGlnAlaHisLeuHisSerLeuCyLeuLeuMetLeuTyriLeuAlaThrGly 20
 DB 88 ATGATCTCTCAGGCCCATCTTCACTCCCTGTGCTTCTTATGCTTATTGGCACTCGA 147
 QY 21 TTTGTLGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
 DB 148 TATGGCCAAAGAGGGGAAGTTTATGTGACCCCTGAAACCCATGACATTTCTATTATGAA 207

QY 41 GlyGlnGlnProSerGlnIleIlePheGlnPheValAspProProAlaValThrPhe 60
 DB 208 GGCAGAAAGACCGATCAATTAATTCAGCTTTAAAGCCATCTCTGCTGTGACTTTT 267
 QY 61 GlnLeuThrGlyGlnThrAspAsnIlePheValIleGlnArgGlnGlyLeuLeuTyriTy 80
 DB 268 GAACCTAAGTGGGAGACAGACAAACATATTTGATGATGAACGGGAGGACTTCTGTATTAC 327
 QY 81 AsnArgAlaLeuAspArgGlnThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 DB 328 AACAGAGCCTTGACAGAGGAAACAAAGATCTACTCAAAATCTCCAGGTTCCAGCCTGAC 387
 QY 101 AlaAsnGlyIleIleValGlnGlnGlnProValProIleThrIleGlnValIleAspIleAsn 120
 DB 388 GCTAATGAAATTAATATGAGAGGCTCAGTCCCTATACCAATGAAAGGACATCAAC 447
 QY 121 AspAsnArgProThrPheLeuGlnSerIleTyriGlnGlySerValArgGlnAsnSerArg 140
 DB 448 GACAAATCGAACCCAGTTTCTCCAGTCAAGTACGAAGGCTCAGTAAAGGCAAACTCTCG 507
 QY 141 ProGlyIleProPheLeuTyriValAlaAlaThrAspLeuAspAspProAlaThrProAsn 160
 DB 508 CCAAGAAAGCCCTTCTTGTATGATCAATGACACAGACCTGATGATCCGGCACTCCCAAT 567
 QY 161 GlnLeuTyriTyriGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyriPhe 180
 DB 568 GGCACGCTTATTAACAGATTCATGATTCACAGCTTCCATATCAACATGCTATGCTTT 627
 QY 181 GlnIleAsnAsnIleThrGlyAlaIleSerLeuThrArgGlnGlySerGlnGlnLeuAsn 200
 DB 628 CAGATCAACAAACAAACGGAGAGCATCTCTTACCAGAGGAGATCTCAGGAATTAATGAAT 687
 QY 201 ProAlaIleAsnProSerTyriAsnLeuValIleSerValIleAspMetGlyGlnSer 220
 DB 688 CCGTCAAAATCCCTTCTTAATATCGATGATCTCAGTAAAGACATGGAGGCCAGAT 747
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGlnAsnIleTyriPhe 240
 DB 748 GAGAAATCTTCAAGTATACGACATCTGTGATATCATATGACAGAGAAATTTGGAAA 807
 QY 241 AlaProIleProValGlnLeuValGlnAsnSerThrAspProIleIleTyriPhe 260
 DB 808 GACCCAAACCTGTGAGATGGGAAACTCACTGATCCACCCCACTCAAAATCACT 867
 QY 261 GlnValIleTyriPheAspProGlyAlaGlnTyriSerLeuValAspIleGlnIlePro 280
 DB 868 CAGGTCCGCTGAAATGATCCCGTGCACAAATTTCTTATTGTTGACAAAGAAAGCTGCCA 927
 QY 281 ArgPheProPheSerIleAspGlnGlnGlnIleTyriValThrGlnProIleAspArg 300
 DB 928 AGATTCCTCAATTTCAATTAACAGAGAGGAAATTTAGTGAATCAGCCCTTGACCA 987
 QY 301 GlnGlnIleAspAlaTyriValPheTyriAlaValAlaIleAspGlnTyriGlyIleProIle 320
 DB 988 GAAAGAAAGATGATATTTTATTAATGATGCAAGATGCAAGATGCAAGATGCAAGAT 1047
 QY 321 SerTyriProIleGlnIleHisValIleValIleAspIleAsnAspAsnProThrTyri 340
 DB 1048 TCATATCCCTCGAATTCATGTAAGTAAAGTAAATTAATTAATGATATCACCCTACATGT 1107
 QY 341 ProSerProValThrValPheGlnValGlnGlnIleAsnGlnIleAsnSerIleGly 360
 DB 1108 CCGTCAACGATTAACGATTTTGAAGTCCAGGAAATGAACGACTGGATTAACGATTCGGG 1167
 QY 361 ThrLeuThrAlaHisAspArgAspGlnGlnIleAsnThrAlaAsnSerPheLeuAsnTyriArg 380
 DB 1168 ACCCTTACTGACATGACAGAGGATGAAGAAATTAATGCAACAGTTTCTTAATCAACAGG 1227
 QY 381 IleValGlnGlnIleProIleLeuProIleAspGlyLeuPheLeuIleGlnIleThrTyriAla 400
 DB 1228 ATTTGAGAACAAATCCCAACTCTCCATGATGACATCTTCTTAATCAAACTATGCT 1287

QY 401 GYMETLEUENLEUVALYSGINSELEUYSLEUVALPHEPHTPROGINTYRASNLEU 420
 DB 1288 GGATGTGTACAGTTAGTAAACAGTCTGTGAAGAGAAAGAACTCTTCTGTAACACTTA 1347
 QY 421 ThrlleluvalSeraplySaspPheleYthrleuCyaspheValGlnleasnValle 440
 DB 1348 ACCATNAGAGGTGTGCAAGAAATTTCAAGCCCTTTGTTTGTGCAAAATCAAGTTAT 1407
 QY 441 AsplleasnApGlnlleProllePheGluYSeraplyrGlyYanleuThrleuAla 460
 DB 1408 GATATCAATATCATGATCCCATCTTGTGAATAATCATATATGAAACCTGACTCTTGT 1467
 QY 461 GlusapPheAnlleGlySerThrleleuThrleleGlnleThrAspAlaAspGluPro 480
 DB 1468 GAAGACACAAACATGSGGTCCACCATTTAAACATCAAGGCCACTGATGCTGATGACCA 1527
 QY 481 PheThrlySerSerlySleleuYrHislelelySglYaspSerGluYargleu 500
 DB 1528 TTTACTGGAGTTTAAATTTCTGTATCATATCAATAAGGAGACATGAGGAGCGCTG 1587
 QY 501 GlyValAspThrAspProHleThrAsnThrGlyYrVallelelySglYProleuAsp 520
 DB 1588 GGGGTGACACAGATCCCATACCAACACCGGATATGTCATTAATTAAGCCCTTGAT 1647
 QY 521 PheGluThrAlaAlaValSerAsnleValPheValAlaGluAsnProGluProleuVal 540
 DB 1648 TTTGAAACAGAGCTGTTTCCACATTTGTTCAAGAGCAAAATCTTGAGCTTGTG 1707
 QY 541 PheGlyVallySlyrAsnAlaSerSerPheAlaYsPheThrleuValleThrAspVal 560
 DB 1708 TTTGGTGTGAAGTACATGCAAGATTTCTTTTGCAGATTCAGCTTATTTGTACAGATGTG 1767
 QY 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaYsValSerGluAspValAla 580
 DB 1768 AATGACACACTCAATTTTCCCAACAGTATTCAGAGGAAAGTCAGTGAAGATGTAGCT 1827
 QY 581 lleGlyThrlySValGlyAsnValThrAlaYsAspProGluYleuAspIleSerTyr 600
 DB 1828 ATAGGCACTAAGTGGGCAATGTGATCTGCCAAGATCCAGAAAGTCTTGACATAGCTAT 1887
 QY 601 SerleuArglyAspThrArglyYrPheleuYsleAspHisValleThrGlyGlnlePhe 620
 DB 1888 TCATGTAGGGGAGACAAAGAGGTGGCTTAATTAATTCACGAGCTGGTGAATCTTT 1947
 QY 621 SerValAlaProleuAspArglyGluAlaGlySerProYrArgValGlnValAlaThr 640
 DB 1948 AGTGTGGTCCATTTGACAGAGAGCCGGAAGTCCATTCGGTACCAAGTGTGCGACA 2007
 QY 641 GluValGlyGlySerSerleuSerSerValSerGluPheHleuMetAspVal 660
 DB 2008 GAAGTGAAGGGGTCTTCTTAACTCTGTGTCAAGTTCCACTGATCTTATGATGTG 2067
 QY 661 AsnAspAsnProProArgleuAlaYsAspYrThrGlyleuPhePheCyHisProleu 680
 DB 2068 AATGACAACTCCCAAGCTAGCAAGAGACTACAGGGCTTCTTCTGCAATCCCTC 2127
 QY 681 SerAlaProGlySerleuIlePheGluAlaThrAspAspAspGlnHisleuPheArgly 700
 DB 2128 AGTCACACCTGGAAGTCTCATTTTCAGGCTACTGATGATGATGACACTTATTTTGGGGT 2187
 QY 701 ProHlePheThrPheSerleuGlySerGlySerleuGlnAsnAspTrpGluValSerlyS 720
 DB 2188 CCCCATTTCATTTTCCCTCGGACGTGAAGCTTACAAACGACTGGGAAGTTTCCAA 2247
 QY 721 lleasnGlyThrHisAlaArgleuSerThrArgHisThrAspPheGluGluAlaYr 740
 DB 2248 ATCATGTACTCATGCCCCGAGCTGTCTACAGGACACAGACTTTGAGAGAGGGCGAT 2307
 QY 741 ValValleuIleArglyleAsnAspGlyGlyArgProProleuGluGlylleValSerleu 760
 DB 2308 GTGGCTTGATCCGCATCAATGATGGGGTGGCGCACTTGGAGAGCATTTGTTCTT 2367
 QY 761 ProValThrPheCySerCyValGluGlySerCyPheArgProAlaGlyHisGlnThr 780

DB 2368 CCAGTTACTCTGCGAGTTGTGTGGAAGAGATGTTTCCGGCCAGAGGTACCAAGACT 2427
 QY 781 GlylleProThrValGlyMetAlaValGlylleleuThrThrleuValleGly 800
 DB 2428 GGGATACCCACGTGGGCAATGAGTGTGATATCTGTACCACTTCTGTGATGTGT 2487
 QY 801 lleleleuAlaValAlaPheIleArglyleYslySaspPlyGlyLysAspAsnValGlu 820
 DB 2488 ATATTTTGACAGTTGTGTTATCCCATTAAGAAAGATTAAGGCAAGATATGTGAA 2547
 QY 821 SerAlaGlnAlaSerGluVallySProleuAspSer 832
 DB 2548 AGTGTCAAGCATCTGAAGTCAACCTCTGAGAAAGC 2583
 RESULT 6
 ABN97312 standard; DNA; 3654 BP.
 ID ABN97312.
 XX AC ABN97312.
 XX AC 13-ANG-2002 (first entry)
 DT XX
 DE XX
 DE Gene #3810 used to diagnose liver cancer.
 KM Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 KM metastatic liver tumor; cyrostatic; expression profile; disease state;
 KM disease progression; drug toxicity; drug efficacy; drug metabolism.
 OS Homo sapiens.
 PN W0200229103-A2.
 PD 11-APR-2002.
 XX PF 02-OCT-2001; 2001WO-US30589.
 PF 02-OCT-2000; 2000US-237054P.
 PR 02-OCT-2000; 2000US-237054P.
 XX (GENE-) GENE LOGIC INC.
 PA Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 PI WPI; 2002-426119/45.
 DR WPI; 2002-426119/45.
 XX PT Diagnosing and detecting the progression of liver cancer.
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample
 XX
 PS Claim 1; SEQ ID NO 3810; 299bp; English.
 XX
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cyostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WPI
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 3654 BP; 1052 A; 804 C; 777 G; 1021 T; 0 other;
 Alignment Scores: 0 Length: 3654
 Pred. No.: 4303.00 Matches: 828
 Score:

Percent Similarity: 99.76%
 Best Local Similarity: 99.52%
 Query Match: 99.58%
 DB: 24
 Gaps: 0

US-10-025-380-1081 (1-832) x ABN97312 (1-3654)

QY 1 MetIleuGlnAlaHisLeuHisSerLeuCySeuLeuMetLeuTyLeuAlaThrGly 20
 DB 93 ATGATACCTCAGGCCCACTTCTCCTGTCCTTATGCTTTATTTGGCACTGGA 152
 QY 21 TyGIglnGlnGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyGlu 40
 DB 153 TATGGCCAGAGGGGGAAGTTTATGTCAGCCCTGAACCCATGACATTTCTATTATGAA 212
 QY 41 GIGlnGlnIupProSerGlnIleIlePheGlnPheLysAlaSerProAlaValThrPhe 60
 DB 213 GGGCCAAAGCCGATCAAAATTAATTCAGTTTAAGGCCAATCCTCGCTGTCATTTT 272
 QY 61 GluLeuThrGlyGlnThrAspAsnIlePheValIleGluArgGlnGlyLeuLeuTyTy 80
 DB 273 GAACTAATCTGGGAGACAGACACATATTTGTGTATAGAGGGGAGACTTCTGTATAC 332
 QY 81 AsnArgAlaLeuAspArgGlnThrArgSerThrHisAsnLeuGlnValAlaLeuAsp 100
 DB 333 AACAGAGCTTGGACAGGGAACAAAGATCTACTCAATCTCCAGTTGACGCCCTGGAC 392
 QY 101 AlaAsnGlyIleIleValGlnGlyProValProIleThrIleGluValLysAspIleAsn 120
 DB 393 GCTAATGGAATTAATAGAGAGGTCAGCTCCCTATACCAATAAAGTGAAGGACATCAAC 452
 QY 121 AspAsnArgProThrPheLeuGlnSerLysTyGIglnGlySerValArgGlnAsnSerArg 140
 DB 453 GACATGAGACCCAGCTTCTCCAGTCAAGTACAAAGGCTCAGTAAGGACAACTCCGC 512
 QY 141 ProGlyLysProPheLeuTyValAsnAlaThrAspLeuAspProAlaThrProAsn 160
 DB 513 CCAAGAAAGCCCTTCTGTATGTCATGCAAGACCAAGACTGATGATCGGACCACTCCAAAT 572
 QY 161 GIGlnLeuTyTyTyGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyPhe 180
 DB 573 GGGCAGCTTTATTAACGATGTCATCAGCTTCCCAATGATCAACAATGTCATGATCTTT 632
 QY 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGlnGlySerGlnGlnLeuAsn 200
 DB 633 CAGATCAACAACAAAGGAGGACATCTCTTACCAGAGGAGATTCAGGAATTCAT 652
 QY 201 ProAlaLysAsnProSerTyAsnLeuValIleSerValLysAspMetGlyGlnSer 220
 DB 693 CTGCTAAGAACTCTTCTTAATCTGATGTCAGTGAAGGACATGGAGGCGCAGAGT 752
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGlnAsnIleTyrPhe 240
 DB 753 GAAATATCTCTTCAAGTATCAACATCTGTGATATATATGACAGGAATATTTGGAAA 812
 QY 241 AlaProLysProValGlnMetValGluAsnSerThrAspProHisProIleLysIleThr 260
 DB 813 GCACCAAAACCTTGAGATGAGTGGGAAAATCTCAACTGATCTCAACCCATCAAAATCACT 872
 QY 261 GlnValArgTyrAsnAspProGlyAlaGlnTySerLeuValAspLysGlnLysLeuPro 280
 DB 873 CAGGTGCGGTGAATGATCCGGTGCACATATCTTATGTCACAAAGGAAGCTGCCA 932
 QY 281 ArgPheProPheSerIleAspGlnGlnGlyAspIleTyTyValThrGlnProLeuAspArg 300
 DB 933 AGATTCCTCATTTCAATGACAGGAAGGAGATATTTACGTGACTCAGCCCTTGACCGCA 992
 QY 301 GluGlnLysAspAlaTyValPheTyValAlaValAlaLysAspGlnTyTyGlyLysProLeu 320
 DB 993 GAAGAAAAGATGCAATATGTTTTTATGCAAGTTGCAGAAAGATGAGTACGAAAACCACTT 1052
 QY 321 SerTyProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProThrTy 340

DB 1053 TCATATCCGCTGGAATTCATGTAAGTAAAGTAAATATATGATATATCCACTACATG 1112
 QY 341 ProSerProValThrValPheGlnValGlnGlnAsnGlnArgLeuGlnLysSerIleGly 360
 DB 1113 CCGTCACACTAACCGTATTTGAGTCCAGGAATGAAGACATGGGTAAACAGATCGGG 1172
 QY 361 ThrLeuThrAlaHisAspArgAspGlnGlnAsnThrAlaAsnSerPheLeuAsnTyArg 380
 DB 1173 ACCCTTACGTCAATGACAGGATGGAAGAAAATACGTGCCAAGAGTTTCTTAACTACAG 1232
 QY 381 IleValGlnGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyVal 400
 DB 1233 ATGTGAGCAAACTCCAAACTCCCATGATGAGTCACTTCCATATCCAACTATGCT 1292
 QY 401 GIMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyAsnLeu 420
 DB 1293 GGAATGTACAGTACCTAAACAGTCTTGAAAGACAGAAATCTCCTCAGTACACTTA 1352
 QY 421 ThrIleGlnValSerAspLysAspPheLysThrLeuCyAspPheValGlnIleAsnValIle 440
 DB 1353 AGGATAGAGGTGCTGACAAAGATTCAGAGCCCTTTGTTGTGCAATCAAGCTTAT 1412
 QY 441 AspIleAsnAspGlnIleProIlePheGlnLysSerAspTyTyValAsnLeuThrLeuAla 460
 DB 1413 GATATCAATGATCAGACCCCATCTTGAATAATCAGATTATGAAAACCTGACTTGCT 1472
 QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGlnPro 480
 DB 1473 GAAGACAAACAAATGAGTCCACATCTTAACATCCAGGCCCACTAATGTATGAGACCA 1532
 QY 481 PheThrGlySerSerLysIleLeuTyHisIleIleLysGlyAspSerGlnGlyArgLeu 500
 DB 1533 TTTACTGGAGGTTCTAAATTTCTGTATCATATATAAAGGAGACAGTGAAGGACCCCTG 1592
 QY 501 GlyValAspThrAspProHisThrAsnThrGlyTyValIleIleLysLysProLeuAsp 520
 DB 1593 GGGGTTGACACAGATCCCATACCAACACCGGATATGCTAATATTAAGACCTCTGAT 1652
 QY 521 PheGlnThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 DB 1653 TTTGAAACAGCAGCTTTTCCACATGTCATGTCMAAGAGAAATCTGAGCTCTAGTG 1712
 QY 541 PheGlyValLysTyAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 DB 1713 TTTGGTGTGAAGTACATGCAATCTTTTGGCAAGTTCAAGCTTATTTGTGACAGATGTG 1772
 QY 561 AsnGlnAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 DB 1773 AATGAAGACACTCAATTTTCCCAACAGTATTCAGAGCGAAAGTCAGTGAGATGAGCT 1832
 QY 581 IleGlyThrLysValGlnValAsnValThrAlaLysAspProGlnGlyLeuAspIleSerTy 600
 DB 1833 ATAGGACATTAAGTGGGCAATGATGATGTCAGAAAGATCCAGAAAGTGTGACATTAACCTAT 1892
 QY 601 SerLeuArgGlyAspThrArgGlyTyPheLysIleAspHisValThrGlyLysIlePhe 620
 DB 1893 TCACTAGGGGAGACACAAAGGTTGGCTTAAATTAATGACACAGTCACTGGTGAATCTTT 1952
 QY 621 SerValAlaProLeuAspArgGlnAlaGlySerProTyArgValGlnValAlaAlaThr 640
 DB 1953 AGTGTGGCTCCATTGGACAGAGAAGCGGAAGTCCATATTCGGGTACAGAGTGGGCCACA 2012
 QY 641 GluValGlyLysSerSerLeuSerSerValSerGlnPheHisLeuIleLeuMetAspVal 660
 DB 2013 GAAGTAGGGGGGTCTTCTTGAAGCTTGTGTAGAGTTCCACTGATCTTATGAGATGAG 2072
 QY 661 AsnAspAsnProProArgLeuAlaLysAspTyTyThrGlyLeuPhePheCyHisProLeu 680
 DB 2073 AATGACAACTCCCGAGGCTAGCCAGAGACTACACGGGCTGTCTTCTTCCATCCCTTC 2132
 QY 681 SerAlaProGlySerLeuIlePheGlnAlaThrAspAspAspGlnHisLeuPheArgGly 700
 DB 2133 AGTGACCTGGAAGTCTCATTTTTCAGGCTACGATGATGATCAGCACTTATTTGGGGGT 2192


```

Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTrpGluValSerIys 720
Db 2193 CCCATTTCATTTTCCTCCGCGAGGAAAGCTGACAAAACGCTGGAGAGTTTCCAAA 2252
Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaTyr 740
Db 2253 ATCATATGTAATCTAGTCCCGACCTGTCTACCGACGACACAGAGTTTGAGAGAGGAGTAT 2312
Qy 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSerLeu 760
Db 2313 GTCCTCTTGATCCCATCAAGATGAGGAGGTCGCGACCCCTTGGAAGCATTTGTTCTTTA 2372
Qy 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGluThr 780
Db 2373 CCACTTACATTTCTTCAGAGTTGTGTGAGAGAGAGTTGTTCCGCGACGACGAGTCCACCAACT 2432
Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
Db 2433 GGGATACCCACGTGGGCGATGCGAGTTGATATCTGCGACCACTTCTGTGATTGGT 2492
Qy 801 IleIleLeuAlaValIlePheIleArgIleIysIysAspIysGlyIysAspAsnValGlu 820
Db 2493 ATATATTTAGCAGTGTGTTTATCCGATTAAGAGATTAAGCAAAAGATTAATGTTGAA 2552
Qy 821 SerAlaGlnAlaSerGluValIysProLeuArgSer 832
Db 2553 AGTCTCAGACATCTGAAGTAAACCTCTGGAAGC 2588

RESULT 7
ABL62484
ID ABL62484 standard; DNA; 3654 BP.
AC ABL62484;
XX
XX
XX 15-MAY-2002 (first entry)
XX
DE Colon adenocarcinoma related gene sequence SEQ ID NO:821.
XX
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancers;
XX cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX gene; ds.
XX
XX Homo sapiens.
XX
XX W0200194629-A2.
XX
XX 13-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US10838.
XX
XX 05-JUN-2000; 2000US-209473P.
XX 05-JUN-2000; 2000US-209531P.
XX 18-SEP-2000; 2000US-233133P.
XX 18-SEP-2000; 2000US-233617P.
XX 20-SEP-2000; 2000US-234009P.
XX 20-SEP-2000; 2000US-234034P.
XX 20-SEP-2000; 2000US-234052P.
XX 22-SEP-2000; 2000US-234509P.
XX 22-SEP-2000; 2000US-234567P.
XX 25-SEP-2000; 2000US-234923P.
XX 25-SEP-2000; 2000US-234924P.
XX 25-SEP-2000; 2000US-235077P.
XX 25-SEP-2000; 2000US-235082P.
XX 25-SEP-2000; 2000US-235134P.
XX 25-SEP-2000; 2000US-235280P.
XX 26-SEP-2000; 2000US-235637P.
XX 26-SEP-2000; 2000US-235638P.
XX 27-SEP-2000; 2000US-235711P.
XX 27-SEP-2000; 2000US-235720P.
XX 27-SEP-2000; 2000US-235840P.
XX 27-SEP-2000; 2000US-235863P.

```

```

PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.

XX
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
XX chemical agent to be tested for anti-neoplastic activity, and
XX determining a change in expression of a gene of a signature gene set
XX
XX Claim 1, SEQ ID 821; 44pp: English.
XX
XX
XX The present invention describes a method (M1) for screening for an
XX anti-neoplastic agent. The method involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening
XX an anti-neoplastic agent, and can be used for producing a product which
XX is the data collected with respect to the anti-neoplastic agent as a
XX result of M1, and the data is sufficient to convey the chemical
XX structure and/or properties of the agent. M1 can be used in the
XX treatment of cancer such as colon, breast, stomach, lung, thyroid,
XX oesophageal, ovarian, kidney, prostate or pancreatic cancer,
XX CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
XX CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
XX carcinoma, papillary carcinoma and Wilm's tumour.
XX
XX
XX SQ Sequence 3654 BP; 1052 A; 804 C; 777 G; 1021 T; 0 other;

Alignment Scores:
Pred. No.: 0
Score: 4303.00
Percent Similarity: 99.76%
Best Local Similarity: 99.52%
Query Match: 99.52%
DB: 24
Gaps: 0

US-10-025-380-1081 (1-832) x ABL62484 (1-3654)
Qy 1 MetIleLeuGlnAlaHisLeuHisSerLeuGlyLeuLeuMetLeuTyrLeuAlaThrGly 20
Db 93 ATGATTAATTCAGGCCCATCTTCACTCCCTGTGCTTCTTATGTTATGGCACTGGA 152
Qy 21 TyrGlyGlnGluGlyIysPheSerGlyProLeuIysProMetThrPheSerIleTyrGlu 40
Db 153 TATGCCCAAGAGGGGAAGTTTACTGGAACCCCTGAACCCATGACATTTTCTATTATGAA 212

```

QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheValAlaSerProAlaValThrPhe 60
 Db 213 GGCCCAAAACCGAGTCAATATATATTCAGTTTAAGGCCAATCTCTCGTGGACCTTTT 272
 QY 61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluValArgGluGlyLeuLeuTyrTyr 80
 Db 273 GAACCTAATCTGGGAGACAGACACATATTTGTGATACAGGAGGAGACTTCGTATATAC 332
 QY 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 Db 333 AACAGAGCCTTGGACAGAGAAACAAGATCTACTACAAATCTCCAGTTTGGAGCCCTGGAC 392
 QY 101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValAspIleAsn 120
 Db 393 GCTAAATGGAATTAATAGGAGGGGTCCAGTCCCTATCACCATTAAGGAAGGACATCAAC 452
 QY 121 AspAsnArgProThrPheLeuGlnSerLeuTyrGluGlySerValArgGlnAsnSerArg 140
 Db 453 GACATCTGACCCAGTTCTTCAGTCAAGTACGAAGGCTCAGTAAGGCAAACTCTGCG 512
 QY 141 ProGlyLysProPheLeuTyrValAlaAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
 Db 513 CAGAGAAAGCCCTTCTTGATGTCAATGCCACAGACCTGGATGATCCGGCACTCCCAAT 572
 QY 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
 Db 573 GGCACACTTATTAACAGATTTGATCAGTCCAGCTTCCCATGATCAACATGTCAATGTACTTT 632
 QY 181 GlnIleAsnAsnLysThrGlyValAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
 Db 633 CAGATCAACAAACAAACGAGGAGCATCTCTTAACCGAGAGGAGATCTCGAAATGGAT 692
 QY 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnGlnSer 220
 Db 693 CCGCTAAGATCTCTTCTTAATCTGTGATCTCAGTGAAGACATGGAGGCCAGAGT 752
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGlnAsnIleTyrLys 240
 Db 753 GAAATATCTTCAAGTATACACATCTGTGATATCATAGTACAGAGAAATTTTGGAAA 812
 QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
 Db 813 GCACCAAAACCTGTGAGATGTGGAACACTCAACGATCTCCACCCCAACAAATCACAT 872
 QY 261 GlnValArgTyrAsnAspProGlyValAlaGlnTyrSerLeuValAspLysGluLysLeuPro 280
 Db 873 CAGGTCCGGTGGATATATCCCGGTGCACATATTTCTTACTTGAACAAAGAAAGCTGCCA 932
 QY 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrValThrGlnProLeuAspArg 300
 Db 933 AGATTCCTCATTTTCAATTTGACAGAGAGGATATTTAGCTGACTCAGGCCCTTGGACCGA 992
 QY 301 GluGlnLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu 320
 Db 993 GAAGAAAGAGATGATATGTTTTTATGCAAGTTCGAAAGATGATGACGAAACACACTT 1052
 QY 321 SerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProProThrCys 340
 Db 1053 TCATATCCCGTGGAAATTCATGTAAGATTAAAGATTATATATATCACTACATATG 1112
 QY 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyLysSerIleGly 360
 Db 1113 CCGTACCAAGTAAACCTATTTGAGGTCCAGAGAAATGAAGACATGGGTAAAGTATCCGG 1172
 QY 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
 Db 1173 ACCCTTAACGACATACAGAGGATGAAGAAATATACGCCAACAGTTTCTTAACCTACAG 1232
 QY 381 IleValGluGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 Db 1233 ATTTGTGAGACAAATCTCCAAACTCTCCATGATGATGACTCTTCTTAATCAAAACATATGCT 1292
 QY 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysLysGlnAspThrProGlnTyrAsnLeu 420

Db 1293 GGAATGTTACAGTTACTTAACAAGCTCTTGAAGAACAGATACTCTCAGTACACTTA 1352
 QY 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
 Db 1353 ACGATAGAGGTGTCTACAAAGATTTTCAAGACCTTTGTTTGTGCAAAATCAACGTTATTT 1412
 QY 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAla 460
 Db 1413 GATATCAATGATAGACCCCATCTTTGAAAAATTCGATATATGGAACCTGACTCTTGCT 1472
 QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db 1473 GAAGACAAACACTTGGTCCACCATCTTAACCATCAAGGCCCTGATGCTGATGAGCCA 1532
 QY 481 PheThrGlySerSerTyrIleLeuTyrHisIleIleLysGlyLysAspArgGluArgLeu 500
 Db 1533 TTTACTGGAGGATTTAAATTTCTGATCATATCATAAAGGAGACAGTAGGAGCCCTTG 1592
 QY 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
 Db 1593 GGGGTGACACAGATCCCATACCAACACCGGATATGTATTAATAAAGCTCTTGAT 1652
 QY 521 PheGluThrAlaAlaLysSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 Db 1653 TTTGAACAGCAGCTGTTTCCAACATGTGTTCAAGAGAAATCTGAGCCTCAATG 1712
 QY 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 Db 1713 TTTGGTGAAGATCAATCAAGTCTTTTGGCAAGTTCAAGTTCATGTTATGTCAGATGTG 1772
 QY 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 Db 1773 AATGAAGCACTCAATTTTCCACACGTAATTCGAAGCAAAAGTCACTGATGATGACT 1832
 QY 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTyr 600
 Db 1833 ATAGGCATTAAGTGGGCAATGACTGCCAAGATCCAGAAAGTCTGGACATTAAGCTAT 1892
 QY 601 SerLeuArgGlyAspThrArgGlyTyrLeuLysIleAspHisValThrGlyLysIlePhe 620
 Db 1893 TCACTGAGGGGAGACACAAAGAGTGGCTTAATAATGGACACAGTCACTGTGATGATCTT 1952
 QY 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaThr 640
 Db 1953 AGTGTGGCTTCCATTGGACAGAGAGCGGAAGTCCATATCGGTACAAAGTGGGCGCAC 2012
 QY 641 GluValGlyLysSerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 Db 2013 GAAGTAAAGGGGTCTTCTTGAAGCTGTGTGACAGATTCACATATCTTATGATATG 2072
 QY 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLysPhePheCysHisProLeu 680
 Db 2073 AATGACAAACCTCCAGAGGTACCAAGGACTACAGGGGCTTGTCTTCCATCCCTTC 2132
 QY 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspGlnHisLeuPheArgGly 700
 Db 2133 AGTGAACCTGGAAGTCTCATTTTCAGGCTACTGATGATGATGACACTTATTTGCGGGT 2192
 QY 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAspTyrGlnValSerLys 720
 Db 2199 CCCCAATTTACATTTTCCCTCGGAGTGAACCTTACAAAGACTGGGAATTTCCAA 2252
 QY 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGlnGluArgAlaTyr 740
 Db 2253 ATCAATGTGACTCAATGCCGCTGCTACAGGACACACAGATTTGAGGAGAGGAGTAT 2312
 QY 741 ValIleLeuIleArgIleAsnAspGlyGlyArgProProLeuGlnGlyIleValSerLeu 760
 Db 2313 GTTCGCTTGATCCGATCAATGATGGGGGTGGCCACTTGGAAAGCAATGTCTTTTA 2372
 QY 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780

Db 2373 CCAGTTACATTCGTCAGTTGTGNGAAGAAAGTTGTTCCGGCAGAGGTCAACAGACT 2432
 QY 781 GYIleProThValGlyMetAlaValGlyIleLeuLeuThrIleLeuValIleGly 800
 Db 2433 GGGATACCCCACTGGGCGATGCGAGTTGTATCTGTCGACCACTTCTGGTGATGGT 2492
 QY 801 ILeIleLeuAlaValAlaPheIleArgIleLeuLeuAspIleGlyLeuAspAlaValGlu 820
 Db 2493 ATATATTTAGCAGTTGTTTATCCGATTAAGAGATTAAGCAAGATTAATTTGAA 2552
 QY 821 SerAlaGlnAlaSerGluValIleProLeuArgSer 832
 Db 2553 AGTGTCAAGCATCTGAAGTCAAACCTCTGGAAGC 2588
 RESULT 8
 ABL67667
 ID ABL67667 standard; DNA; 3654 BP.
 XX
 AC ABL67667;
 DT 15-MAY-2002 (first entry)
 XX
 DE Oesophagus cancer related gene sequence SEQ ID NO: 6004.
 XX
 KW Human, cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW Stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW Cytosolastic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN M0200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001MO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-231133P.
 PR 18-SEP-2000; 2000US-231617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR

PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVALON) AVALON PHARM.
 XX
 XX Young PE, Augustus M, Carter KC, Edner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 DR WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 PS Claim 1; SEQ ID 6004; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.
 XX
 SQ Sequence 3654 BP; 1052 A; 804 C; 777 G; 1021 T; 0 other;
 Alignment Scores:
 Pred. No.: 0 Length: 3654
 Score: 4303.00 Matches: 828
 Percent Similarity: 99.76% Conservative: 2
 Best Local Similarity: 99.52% Mismatches: 0
 Query Match: 99.58% Indels: 0
 DB: 24 Gaps: 0
 US-10-025-380-1081 (1-832) x ABL67667 (1-3654)
 QY 1 MetIleuGlnAlaHisIleuHisSerLeuCySLeuLeuMetLeuTyrLeuAlaThrGly 20
 Db 93 ATATATCTTCAGGCGCCATCTTCACTCCCTGTGTCTTATGTTATTTGGCACTGGA 152
 QY 21 TyrGlyGlnGluPheSerGlnIleIlePheGlnPheLeuAlaAspProPheAlaValThrPhe 40
 Db 153 TATGCCCAAGAGGGAGAGTTAGTGACCCCTGAACCATGACATTTCTATTATGAA 212
 QY 41 GlyGlnGluPheSerGlnIleIlePheGlnPheLeuAlaAspProPheAlaValThrPhe 60
 Db 213 GGCCCAAGAACCGAGCAATTAATATTCAGTTTAAGGCCATCTCTGCTGACCTTT 272
 QY 61 GlnLeuThrGlyGlnThrAspAlaIlePheValIleGlnArgGlyGlyLeuLeuTyrTyr 80
 Db 273 GAACCTAACCTGGGAGAGCAACATATTTGTATGAAAGGAGGACCTTCTGTATTAC 332
 QY 81 AsnArgAlaLeuAspArgGlnThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 Db 333 AACAGACCTTTGACAGGAGAAACAAGATCTACACAAATCTCAGGTGACCCCTGGAC 392
 QY 101 AlaAsnGlyIleIleValaGlyGlyProValProIleThrIleGluValIleAspIleAsn 120
 Db 393 GCTAATGGAATTAATAGAGGGGTCCAGTCCCTATCAACCATTAAGTAAGGACATCAAC 452

QY 121 AspaanaArgProThrPheLeuGlnSerIysTyrgIuGlySerValArgGlnAasnSerArg 140
Db 453 GACAAATGACCCAGCTTCTCCAGTCAAGTAAGAGGCTCAGTAAGGCAAACTCTCGC 512
QY 141 ProGlyIysProPheLeuTyrrValAasnAlaThrAspLeuAspPheProAlaThrProasn 160
Db 513 CCAGAAAGGCCCTTCTGTATGTAATGCCACAGACTGATGATCCGGCCACCTCCCAAT 572
QY 161 GlyGlnLeuTyrrTyrgIlnIleValIleGlnLeuProMetIleAasnValMetTyrrPhe 180
Db 573 GGCCAGCTTTATTCAGATGTCATCCAGCTTCCATGATCAAAATGTCATGACTT 632
QY 181 GlnIleAasnIlysthrGlyAlaIleSerLeuThrArgGlyIuGlySerGlnIuLeuAasn 200
Db 633 CAGATCAACAACAAGGCGGATCTCTCTTACCCGAGGAGATCTCAGGAATTGAT 692
QY 201 ProAlaIysAsnProSerTyrrAasnLeuValIleSerValIysAspMetGlyGlnIser 220
Db 693 CCTGCTAAGAACTCTCTTAATCTGTGATCTCAGTGAAGGACATGGGAGCCAGAGT 752
QY 221 GluAasnSerPheSerAspThrThrSerValAspIleIleValIleThrgIuAasnIleTrpIys 240
Db 753 GAGAAATCTTCAAGTAACCAATCTGTGATATCATGACAGAGAAATATTGAAAA 812
QY 241 AlaProIysProValGluMetValGluAasnSerThrAspProHisProIleIysIleThr 260
Db 813 GCACCAAAACCTGTGAGATGTGAAAACTCAACTATCTCACTCCCATCAAAATCACT 872
QY 261 GlnValAlaGTrpAasnAspProGlyAlaGlnTyrrSerLeuValAspIlyGlyIysLeuPro 280
Db 873 CAGGTGGGTGGAATGATCCCGGTGCACAATATCTCTAGTTGCAAAAGAGAGCTCCA 932
QY 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrrValIleThrgIuProLeuAspArg 300
Db 933 AGATTTCCCATTTTCAATTGACCAAGAGAGAAATTAATTCGACTCAGCCCTTGAGACGA 992
QY 301 GluGlnIlyAspAlaTyrrValPheTyrrAlaValAlaIysAspGlyTyrrGlyIysProLeu 320
Db 993 GAAGAAAGAGATGATATGTTTTTATGCAAGTTGCAAGGATGATGCGGAAAAACCACTT 1052
QY 321 SerTyrrProLeuGlnIleHisValIysValIysAspIleAasnAspAsnProProThrCys 340
Db 1053 TCATATCCGCTGGAATTCATGTAAAGTTAAAGATTAATGATATTCACCACTCACTGAT 1112
QY 341 ProSerProValIleThrValPheGlnValGlnGluAasnGlnArgLeuGlyAasnSerIleGly 360
Db 1113 CCGTCAACAGTAACCGTAATTGAGGTCCAGAGATGAACGACTGGGTAAACAGATCGGG 1172
QY 361 ThrLeuThrAlaHisAspArgAspGluGluAasnThrAlaAasnSerPheLeuAasnTyrrArg 380
Db 1173 ACCCTTACTGCATGACAGGATGAGAAAGAAATTAATCTCCAAAGCTTTCTTAACTACAGG 1232
QY 381 IleValGlnGlnThrProIysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrrAla 400
Db 1233 ATGTGTGAGCAAACTCCCAAACTTCCATGATGATGATCTTCTTAATCCAAACTATGCT 1292
QY 401 GlyMetLeuGlnIleuAlaIysGlnSerLeuIlySerGlnAasnThrProGlnTyrrAasnLeu 420
Db 1293 GGAATGTACAGTATGCTAAACAGTCTTGAAGAAGAAAGATATCTCTCGATCAACTTA 1352
QY 421 ThrIleGlnValSerAspIysAspPheIysThrLeuCysPheValGlnIleAasnValIle 440
Db 1353 ACAGTAGAGGTGTCTGCACAAAGATTTCAGACCTTTGTTTGTGCAAAATCAACGTTATT 1412
QY 441 AspIleAasnAspGlnIleProIlePheGlnIlySerAspTyrrGlyAasnLeuThrLeuAla 460
Db 1413 GATATCATATGATCAACCCCATCTTGAAGAAATCAATTAATGAGAAACCTGACCTTGCT 1472
QY 461 GluAspThrAasnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
Db 1473 GAAGACACAAACATTGGGTCCACCATCTTAACATCAAGGCCACTGATGCTGATGAGCCA 1532

QY 481 PheThrGlySerSerIysIleLeuTyrrHisIleIleIysGlyAspSerGlnGlyArgLeu 500
Db 1533 TTTACTGGGAGATTCTAAATTTCTGTATCATATCATAAAGGAGACAGTGAAGGACGCTG 1592
QY 501 GlyValAspThrAspProHisThrAsnThrGlyTyrrValIleIleIysIysProLeuAsp 520
Db 1593 GGGGTGACACAAATCCCAATCAACACCGGATATGCTAAATTAATAAAGCCCTCTTGAT 1652
QY 521 PheGlnThrAlaAlaIysAsnIleValPheIysAlaGluAasnProGluProLeuVal 540
Db 1653 TTTGAAACAGCAGCTGTTTCCAAATTTGTTTCAAGACGAAATCTGAGCCTTATGTC 1712
QY 541 PheGlyValIysTyrrAasnAlaSerSerPheAlaIysPheThrLeuIleValIleAspVal 560
Db 1713 TTTGTGTGAAGTACATGATCAATGTTTGTGCAAGTTCACGCTTAATGTGACAGATGTC 1772
QY 561 AsnGlnAlaProGlnPheSerGlnHisValPheGlnAlaIysValSerGluAspValAla 580
Db 1773 AATGAAGCACCTCAATTTTCCCAACAGTAATCCAAAGTCGAAAGTCAGTGAAGTGTAGCT 1832
QY 581 IleGlyThrIysValGlyAasnValIleThrAlaIysAspProGluIleuAspIleSerTyrr 600
Db 1833 ATAGGACTAAAGTGGCAATGATGACTGCCAAGGATCCAGAGGTCGACATPAAGCTAT 1892
QY 601 SerLeuArgGlyAspThrArgGlyTyrrLeuIysIleAspHisValIleThrgIyGlnIlePhe 620
Db 1893 TCACCTGAGGGGAGACCAAGAGTGGCTTAATAATGACACAGCTGATGATGATCTTT 1952
QY 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrrArgValGlnValIleAlaThr 640
Db 1953 AGTGTGCTCCATTTGACAGAGAAAGCCGGAAGTCATATCGGTACAGAGTGGGCCACA 2012
QY 641 GluValGlyIysSerSerIysSerSerValIserGlnPheHisIleuIleLeuMetAspVal 660
Db 2013 GAGATAGGGGGGCTTCTTCTGACCTGTGTCAAGTCTCAGCTTCACTTAATGAGATGTG 2072
QY 661 AsnAspAsnProProArgLeuAlaIysAspTyrrThrgIyLeuPhePheCysHisProLeu 680
Db 2073 AATGACAAACCTCCAGGGGTAGCCAAAGACTACACGGGGTGTGTTCTTGTCCATCCCTC 2132
QY 681 SerAlaProGlySerLeuIlePheGlnAlaThrAspAspAspGlnHisIleuPheArgGly 700
Db 2133 AGTGACCTCGAAGTCTCATTTTCGAGGCTACGATGATGATCAGACTTAATTCGGGGT 2192
QY 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGluAasnAspTrpGluValIleSerIys 720
Db 2193 CCCCATTTTACATTTTCCCTCGGACGTGAAGCTTCAAAACGACTGGGAAGTTTCCAA 2252
QY 721 IleAasnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGlnGluArgAlaTyrr 740
Db 2253 ATCAATGTAATCTATATCCGACCTGTCTACCAAGCAACAGAGTTTAGAGAGAGAGTAT 2312
QY 741 ValValleuIleArgIleAasnAspGlyIysArgProProLeuGlnGlyIleValIleSerLeu 760
Db 2313 GTGCTCTTGAATCCGATCAATATGAGGGGTGGCCACCTTGGAAAGCATTTGTTCTTAA 2372
QY 761 ProValIlePheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
Db 2373 CCAATTACATCTGCAAGTGTGTGGAAGAAAGTTGTTCCGGGACAGAGTCCAGACT 2432
QY 781 GlyIleProThrValGlyMetAlaValIleGlyIleLeuLeuThrThrLeuLeuValIleGly 800
Db 2433 GGGATACCACTGTGGGAGATGGAGATGGATATCTGTACACACCTTCTGTGATTTGTG 2492
QY 801 IleIleLeuAlaValIlePheIleArgIleIysIysAspIysGlyIysAspAasnValGlu 820
Db 2493 ATAAATTTAGCAGTTGTGTTATCCCATAAAGAAAGATPAAGGCAAGATATATGTTGAA 2552
QY 821 SerAlaGlnAlaSerGluValIysProLeuArgSer 832
Db 2553 AGTGTCAAGCATCTGAAGTCAAACTCTGAAGAGC 2588

RESULT 9

AAH98190
 ID AAH98190 standard; cDNA; 2820 BP.
 AC AAH98190;
 DT 12-OCT-2001 (first entry)
 XX
 DE Human EST-derived coding sequence SEQ ID NO: 47.
 XX
 KM Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KM tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KM diagnostics; forensic test; gene mapping; genetic disorder;
 KM biodiversity; gene therapy; nutrition; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200154477-A2.
 PD 02-AUG-2001.
 XX
 PF 25-JAN-2001; 2001WO-US02687.
 XX
 PR 25-JAN-2000; 2000US-0491404.
 PR 17-JUL-2000; 2000US-0617746.
 PR 03-AUG-2000; 2000US-0631451.
 PR 15-SEP-2000; 2000US-0663870.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 PI Cao Y, Dmanac RA, Zhang J, Werhman T;
 XX
 DR WPI: 2001-476164/51.
 DR P-PSDB; AAM23531.
 XX
 PT Isolated polypeptide for treatment of diseases, diagnostics, raising
 PT antibodies and research use -
 XX
 PS Claim 1; Page 213-214; 1275pp; English.
 XX
 CC The present invention provides the protein and coding sequences of novel
 CC proteins from a variety of organisms, including human, dog, cat, horse,
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a cDNA
 CC of the invention.
 CC
 SQ Sequence 2820 BP; 593 A; 867 C; 821 G; 539 T; 0 other;
 CC
 Alignment Scores:
 Pred. No.: 1.19e-70 Length: 2820
 Score: 922.50 Matches: 237
 Percent Similarity: 51.02% Conservative: 163
 Best Local Similarity: 30.23% Mismatches: 353
 Query Match: 21.35% Indels: 31
 DB: 22 Gaps: 19
 US-10-025-380-1081 (1-832) x AAH98190 (1-2820)
 QY 62 LeuHrGlyGlu-----ThrAspAnIlePheValIleGluArgGlu---Gly 76
 DB 220 CTGTCAGGGGACTCAGGCAAGGCACTGAGGCCCATTTGCTATGATCCAGATTCTGGC 279
 QY 77 LeuLeuTyTyrAsnArgAlaLeuAspArgGluThrArgSerThrAsnLeuGlnVal 96
 DB 280 TTCTGCTGCTGACAGGGCCCTGGACCGAGAGAGAGAGAGTACCTACAGGTC 339
 QY 97 AlaAlaLeuAspAlaAsnGlyTleValGluGlyProValProIleThrIleGluVal 116
 DB 340 ACCCTGAGATGACAGATGACATGCTTGTGGGGTCCACAGCCTGTGCTTGCACGTG 399

QY 117 LysAspIleAsnAspAsnArgProThrPheLeuGlnSerTyTyrGluGlySerValArg 136
 DB 400 AAGGATGAGATATACAGAGTGGCCCATTTCTCTCAAGCATCTACAGAGTGGAGTGGAGC 459
 QY 137 GlnAsnSerArgProGlyTyPProPheLeuTyTyrValAsnAlaThrAspLeuAspPro 156
 DB 460 CGGGATACCAAGGCTGGCATCCCTTCCTTGGAGCTTCACAGCGGAGTGAAGCA 519
 QY 157 AlaThrProAsnGlyIleLeuTyTyrGlnIleValIleGlnLeuProMetIleAsnAsn 176
 DB 520 GGCACAGCCAACTCGGATCTTCATTCACATCTCAAGCTCAAGCTCCAGCCAGCTTCC 579
 QY 177 ValMetTyPheGlnIleAsnAsnTyTrnGlyAlaIleSerLeuThrArgGluGlySer 196
 DB 580 CCAGACATGTTCCAGCTGGAGCTGGAGCTGGGGGCTCTGGCCCTCAGCCCAAGGGAGC 639
 QY 197 GlnGluLeuAsnProAlaTyAsnProSerTyTyrAsnLeuValIleSerValAspMet 216
 DB 640 ACCAGCTTGAACCAACCCCTGGAGAGAGACTTACAGCTGTGTACAGGTCAAGACATG 699
 QY 217 GlyGlyGlnSerGluAsnSerPheSerAspThrThrSerValAspIleIleValThrGlu 236
 DB 700 GTGACCAAGGCC--TCAGGCCACAGGCCACTGTCACCTGGAGTCTTCATCATAGAG 756
 QY 237 AsnIleTyPylAspAlaProLysProValGluMetValGluAsnSerThrAspProHisPro 256
 DB 757 AGCAGCTGGGTGTCCTAGAGCTTATCCAGCTGAGAGATCTCAAGTCTATACCCG 816
 QY 257 IleValIleThrGlnValArgTrpAsnAspProGlyAlaGlnTySerLeuValAspLys 276
 DB 817 CACCACATGGCCCAAGGATACAGTGGGGTGTAGTGTGACATCATCAGCTG----- 867
 QY 277 GlnLysLeuProArgPheProPheSerIleAspGlnGluGlyAspIleTyTyrValThrGln 296
 DB 868 GAGAGCCATCCCGGGGACCTTTGAAGTAATCAGAGGGAACCTTCTGACGACAGA 927
 QY 297 ProLeuAspArgGluGluLysAspAlaTyTyrValPheTyTyrAlaValAlaLysAspGluTyTyr 316
 DB 928 GAGCTGGACAGAGAAGCGGCTGAGTACCTCTCCAGGTGGGGCTCGAATTCAT 987
 QY 317 GlyLysProLeuSerTyTyrProLeuGlnIleIleValLysValLysAspIleAsnAspAsn 336
 DB 988 GGGAGAGACTATCCCGGCCCTCTGAGCTGCAGCTGTGTGTAGTGAATGAGATGACAC 1047
 QY 337 ProProThrCysProSerProValThrValPheGluValGlnGluAsnGluArgLeuGly 356
 DB 1048 GTGCTATCTGCTCCCTCCCTCCCTGACCCCAAGTACGATCTTACCTCAGTCCACAGGT 1107
 QY 357 AsnSerIleGlyThrLeuThrAlaAsnAspArgAspGluGluAsnThrAlaAsnSerPhe 376
 DB 1108 ACTGAGTGTACTGACTGTCTGACAGAGATGAGATGCCCGCGGCTCCCATTTCCAC 1167
 QY 377 LeuAsnTyTyrArgIleValGluGlnThrProLysLeuProMetAspGly--LeuPheLeu 395
 DB 1168 GTTGTGATACGCTCTGAGCCCTGAGCCTGAGAGATGGGTGAAGGGAGAGGCTTCCAG 1227
 QY 396 IleGlnThrTyTyrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLysLys--GlnAsp 414
 DB 1228 GTGAGCCCACTTCAGAGAGTGTGACGCTGGGGGTGCTCCCATCTCCAGAGAGCCAGAAC 1287
 QY 415 ThrProGlnTyTyrAsnLeuThrIleGluValSerAspLysAsp-----PheTyThrLeu 432
 DB 1288 ATCTGCTTGTGCTGTGCTGCGCATGAGCTGGAGAGGGGGGGCTTCACAGAGACG 1347
 QY 433 CysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLysSer 452
 DB 1348 TGTGAAGTCAAGATCCAGTCAAGATATCAAGATCAAGCCCTGTGATTCATCATCTCC 1407
 QY 453 AspTyTyrGlyAsnLeuThrIleValGluAspThrAsnIleGlySerThrIleLeuThrIle 472
 DB 1408 CAGATTGGGCTTAAAGCTCTCTGAGAGATGAGAGCCGGGAGCTGTGGGACATCTTA 1467
 QY 473 GlnAlaThrAspAlaAsp--GluProPheThrGlySerSerTyIleLeu--TyTrHis 490


```

Db      281 TTCTGCTGGACCAAGGCGCTGACCCAGAGAGACGAGGAGCTACAGCTCACAGTC 340
Qy      97 AlaAlaLeuAapAlaAnglyLeIleValGluGlyProValProIleThrIleGluVal 116
Db      341 ACCGTGAGATGACGATGATGATGCTTGTGGGGTCCACAGCTGTGCTTGACAGTGTG 400
Qy      117 LysAspIleAsnAapAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAa 136
Db      401 AAGATAGAGATGACGAGTGGCTGCTTCTCAAGCATCTACAGCTGAGCTGAGTGGTACG 460
Qy      137 GlnAsnSerArgProGlyLysProPheLeuValAsnAlaThrAspLeuAspAspPro 156
Db      461 CGGGGTACAGAGCTGGGATCCCTTCTCTCTTCTGAGCTTCAAGCCGAGATGAGACA 520
Qy      157 AlaThrProAanglyGlnLeuValLeuValIleGlnLeuProMetIleAsnAsn 176
Db      521 GGCACACCAACTCGGATCTTCGATTCACATCTGAGCCAGGCTCCAGGCCAGCTTCC 580
Qy      177 ValMetValPheGlnIleAsnAsnLeuThrGlyAlaIleSerLeuThrArgGlySer 196
Db      581 CCAGACATGTTCCAGCTGAGCTCTGGCTGGGGCTCTGGCTTCAAGCCCAAGGGAGGC 640
Qy      197 GlnIleLeuAapProAlaLysAsnProSerValAsnLeuValIleSerValLysAspMet 216
Db      641 ACCAGCTTGACACAGCCCTGGAGAGACCTACAGCTGTGTGACAGCTCAAGCATG 700
Qy      217 GlyGlyGlnSerGluAaSerPheSerAspThrThrSerValAspIleIleValThrGlu 236
Db      701 GGTACACAGGCC--TCAGGCCACACAGGCCACCTGCCAGGAGCTTCATCATGAG 757
Qy      237 AsnIleThrLysAlaProLysProValGluMetValGluAaSerThrAspProHisPro 256
Db      758 AGCACCTGGGTGCTCTAGAGCTTACCACTGGGAGAGATCTCAAGTCTTACCCG 817
Qy      257 IleValIleThrGlnValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLys 276
Db      818 CACACACATGGCCACAGTACACTGAGATGGGGTATGTCACATTCACCTG----- 868
Qy      277 GlnLysLeuProArgPheProPheSerIleAspGlnGlyAspIleTyrValThrGln 296
Db      869 GAGAGCCATCCCGGGGACCTTGTGAAGTGAATGACAGAGGAAACCTTCAAGTACACGA 928
Qy      297 ProLeuAapArgGluGlnLysAspAlaTyrValThrValAlaValAlaAspGluTyr 316
Db      929 GAGCTGGACAGAGGCCAGGCTGAGTACCTGCTCCAGGCTCGGGCTCAGAAATTCCTAT 988
Qy      317 GlnLysProLeuSerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsn 336
Db      989 GGGAGAGACTATGGGGCCCTCTGGAGTGCACGCTGCTGATGATGAGATGAGATGACAC 1048
Qy      337 ProProThrCysProSerProValThrValPheGluValGlnGluAaGluGly 356
Db      1049 GTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1108
Qy      357 AsnSerIleGlyThrLeuThrAlaHisAspArgGluGluAaThrAlaAsnSerPhe 376
Db      1109 ACTGAGTGAAGTACATCTGACAGAGAGATGAGATGAGTGGCTGCTGCTGCTGCTGCTGCT 1168
Qy      377 LeuAsnTyrArgIleValGlnIleThrProLysLeuProMetAspGlyLysLeuPheLeu 395
Db      1169 GTTGTGATACAGCTCTGAGCCCTGAGCTGAGATGGGGTGAAGAGGGAGAGCTTCCAG 1228
Qy      396 IleGlnThrTyrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLysValLysGlnAsp 414
Db      1229 GTGAGACCCCACTGACGAGTGTGACGCTGGGGGTCTCCATCCGAGGAGGAGGAGGAGGAG 1288
Qy      415 ThrProGlnTyrAsnLeuThrIleGluValSerAspLysAsp-----PheLysThrLeu 432
Db      1289 ATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1348
Qy      433 CysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLysSer 452

```

```

Db      1349 TGTGATCGAAGTCCAGTCAAGATATCAATGATACAGCCCTGAGTTCATCACTTCC 1408
Qy      453 AspTyrGlyAsnLeuThrIleValAspIleAsnIleGlySerThrIleLeuThrIle 472
Db      1409 CAGATTGGGCTTAAAGCTCTCCCTGAGAGTGGAGCCCGGAGCTTGGTGGCATCTTA 1468
Qy      473 GlnAlaThrAspAlaAsp---GluProPheThrGlySerSerLysIleLeu---TyrHis 490
Db      1469 ACAGCATTAATCTCACTCCAGGCC-----GCCCTCCGCTCATGATTTTGGC 1519
Qy      491 IleIleLysGlyAspSerGluGlyArgLeuGlyValAspThrAspProHisThrAsnThr 510
Db      1520 ATTGAGAGGGAGACACAGAGGGACTTTGGCTGCTGATTTGGAGCA-----GACTTC 1573
Qy      511 GlyTyrValIleIleLys-----LysProLeuAspPheGlnThrAlaAlaValSerAsn 528
Db      1574 GGGCATTTAGACTCAAGCTTGCAGAAACCTAAGTTAGAGGAGCTCCAACTCATGAG 1633
Qy      529 IleValPheValAlaGluAaProGluProLeuValPheGlyValLysTyrAsnAlaSer 548
Db      1634 GTGTGTGTGTGTGTGACAGTGTGGGAGCTGTG---GGGCCAGGCCAGGCCCTGGA 1690
Qy      549 SerPheAlaLysPheThrLeuIleValThrAspValAsnGluAlaProGlnPheSerGln 568
Db      1691 GCCACGGCCACGGGTGACTGTGCTAGTGAAGAGATGATGCCACCCCAAGTTGACACAG 1750
Qy      569 HisValPheGlnAlaLysValSerGluAaProValAlaIleGlyThrLysValGluAaVal 588
Db      1751 GAGAGCTACAGGACGAGTGTCCCATACAGTCCCAAGCCGCGCTTTCCTGCTGACCATC 1810
Qy      589 ThrAlaLysAspProGluGlyLeuAspIleSerTyrSerLeuArgLysThrArgLys 608
Db      1811 CAGCCCTCCAGACCCCATACAGCCGACACCTCAGGTTCTCCCTAATCATGATCAGAGGGC 1870
Qy      609 TyrLeuLysIleAspHisValThrGlyLysIlePheSerValAlaProLeuAsp---Arg 627
Db      1871 TGCTCTGATGAGAAATCTCCGGGAGGTGCACACCCCGACGCTCCCTGCGAGGGCGCC 1930
Qy      628 GlnAlaGlySerProTyrArgValGlnValAlaIleThrGluValGlyGlySerSerLeu 647
Db      1931 CACCTGGGAGACCTTACACGCTGCTTGGAGGCCAGAGATCAAGATGACCGGAGACTG 1990
Qy      648 SerSerValSerGluPheHisLeuIleMetAspValAsnAspAspProArgLysLeu 667
Db      1991 AGGCTTGTGACACCCCTGGATGATCACTTCAAGGCCCTGCTGCCAGCCGTGACT 2050
Qy      668 AlaLysAspTyrThrGlyLeuPhePheCysHisProLeuSerAlaProGlySerLeuIle 687
Db      2051 CTGCCCCCTGCTGCTCCCAATACCTTGCACACCCCGCCAGACCATGCTGATCTGTG 2110
Qy      688 Phe---GluAlaThrAspAspArgGlnHisLeuPheArgGlyProHisPheThrPheSer 706
Db      2111 AGTGAACCCAGCAAGACCCCGATCTGGCCAGTGGGACAGCTGCC---TACAGCTTACAC 2167
Qy      707 LeuGly---SerGlySerLeuGlnAsnAspTyrGluValSerLysIleAsnGlyThrHis 725
Db      2168 CTGTGCTCCAAACCCACGCTGACAGGAGATTTGGCGCTCAGACTCAGATGTTGCTCAT 2227
Qy      726 AlaArgLeuSerThrArgHisThrAspPheGlnGluArgAlaTyrValValLeuIleArg 745
Db      2228 GCTTACTTCACTTGGCGCTGATGGTGGAGCCACGTTGACACATTAATCCCGCTGTG 2287
Qy      746 IleAsnAspGlyLysArgProProLeuGlnGlyIleValSerLeuProValThrPheCys 765
Db      2288 GTGAGCCCAATGCCAG-----ATGTGCACTCTGTGTGAGATGATGCTGTGT 2338
Qy      766 SerCys---ValGlnGlySerCysPheArgProAlaGlyHisGlnThrGlyIleProThr 784
Db      2339 CGCTGCAACTGAGAGGGAGCTGATGCGAGAGTGGGGCCGATGAAGGCAATGCCAGC 2398
Qy      785 ValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGlyIleIleValAla 804
Db      2399 AAGCTGTGGAGTGGGATCTTTAGGACACCTGTGATGACATTAAGAAATCTTCATC 2458

```

QY 805 ValValhelaIeaIgiIeIyIaApIySgIyIySaPaenValgiuSerAlaGlnAla 824
Db 2459 CTCATTTCACCCACTGCACATGTCACAGGAAGAGCCCGATCACAACGACGACAGC 2518
QY 825 SerGIuValIyS 828
Db 2519 GTGCCCTGAAG 2530

RESULT 11
AAA37087
ID AAA37087 standard; cDNA; 2848 BP.
AC AAA37087;
XX
XX
DT 08-AUG-2000 (first entry)
XX
DE Human PRO1340 (UN0695) cDNA sequence SEQ ID NO:228.
XX
XX Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening;
XX ss.
XX Homo sapiens.
XX
XX MO200012708-A2.
XX
XX 09-MAR-2000.
XX
XX 01-SEP-1999; 99WC-US20111.
XX
XX 01-SEP-1998; 98US-0098716.
PR 01-SEP-1998; 98US-0098749.
PR 01-SEP-1998; 98US-0098750.
PR 02-SEP-1998; 98US-0098803.
PR 02-SEP-1998; 98US-0098821.
PR 02-SEP-1998; 98US-0098843.
PR 03-SEP-1998; 98US-0099536.
PR 03-SEP-1998; 98US-0099596.
PR 03-SEP-1998; 98US-0099598.
PR 03-SEP-1998; 98US-0099602.
PR 03-SEP-1998; 98US-0099642.
PR 10-SEP-1998; 98US-0099741.
PR 10-SEP-1998; 98US-0099754.
PR 10-SEP-1998; 98US-0099763.
PR 10-SEP-1998; 98US-0099792.
PR 10-SEP-1998; 98US-0099808.
PR 10-SEP-1998; 98US-0099812.
PR 10-SEP-1998; 98US-0099815.
PR 10-SEP-1998; 98US-0099816.
PR 15-SEP-1998; 98US-0100385.
PR 15-SEP-1998; 98US-0100388.
PR 15-SEP-1998; 98US-0100390.
PR 16-SEP-1998; 98US-0100584.
PR 16-SEP-1998; 98US-0100627.
PR 16-SEP-1998; 98US-0100661.
PR 16-SEP-1998; 98US-0100662.
PR 16-SEP-1998; 98US-0100664.
PR 16-SEP-1998; 98US-0100665.
PR 17-SEP-1998; 98US-0100683.
PR 17-SEP-1998; 98US-0100710.
PR 17-SEP-1998; 98US-0100711.
PR 17-SEP-1998; 98US-0100919.
PR 17-SEP-1998; 98US-0100930.
PR 18-SEP-1998; 98US-0100848.
PR 18-SEP-1998; 98US-0100849.
PR 18-SEP-1998; 98US-0101014.
PR 18-SEP-1998; 98US-0101068.
PR 18-SEP-1998; 98US-0101071.
PR 22-SEP-1998; 98US-0101279.
PR 23-SEP-1998; 98US-0101471.
PR 23-SEP-1998; 98US-0101472.
PR 23-SEP-1998; 98US-0101474.

PR 23-SEP-1998; 98US-0101475.
PR 23-SEP-1998; 98US-0101476.
PR 23-SEP-1998; 98US-0101477.
PR 23-SEP-1998; 98US-0101479.
PR 24-SEP-1998; 98US-0101738.
PR 24-SEP-1998; 98US-0101741.
PR 24-SEP-1998; 98US-0101743.
PR 24-SEP-1998; 98US-0101915.
PR 24-SEP-1998; 98US-0101916.
PR 24-SEP-1998; 98US-0102207.
PR 24-SEP-1998; 98US-0102240.
PR 29-SEP-1998; 98US-0102307.
PR 29-SEP-1998; 98US-0102310.
PR 29-SEP-1998; 98US-0102330.
PR 30-SEP-1998; 98US-0102331.
PR 30-SEP-1998; 98US-0102484.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102570.
PR 30-SEP-1998; 98US-0102571.
PR 01-OCT-1998; 98US-0102684.
PR 01-OCT-1998; 98US-0102687.
PR 02-OCT-1998; 98US-0102965.
PR 06-OCT-1998; 98US-0103258.
PR 06-OCT-1998; 98US-0103449.
PR 07-OCT-1998; 98US-0103314.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103395.
PR 07-OCT-1998; 98US-0103396.
PR 07-OCT-1998; 98US-0103401.
PR 08-OCT-1998; 98US-0103633.
PR 08-OCT-1998; 98US-0103679.
PR 08-OCT-1998; 98US-0103711.
PR 14-OCT-1998; 98US-0104257.
PR 20-OCT-1998; 98US-0104987.
PR 20-OCT-1998; 98US-0105000.
PR 20-OCT-1998; 98US-0105002.
PR 21-OCT-1998; 98US-0105104.
PR 22-OCT-1998; 98US-0105169.
PR 22-OCT-1998; 98US-0105266.
PR 26-OCT-1998; 98US-0105593.
PR 26-OCT-1998; 98US-0105594.
PR 27-OCT-1998; 98US-0105807.
PR 27-OCT-1998; 98US-0105881.
PR 27-OCT-1998; 98US-0105882.
PR 27-OCT-1998; 98US-0106062.
PR 28-OCT-1998; 98US-0106023.
PR 28-OCT-1998; 98US-0106029.
PR 28-OCT-1998; 98US-0106030.
PR 28-OCT-1998; 98US-0106032.
PR 28-OCT-1998; 98US-0106033.
PR 28-OCT-1998; 98US-0106178.
PR 29-OCT-1998; 98US-0106248.
PR 29-OCT-1998; 98US-0106384.
PR 29-OCT-1998; 98US-0106500.
PR 30-OCT-1998; 98US-0106544.
PR 03-NOV-1998; 98US-0106556.
PR 03-NOV-1998; 98US-0106502.
PR 03-NOV-1998; 98US-0106905.
PR 03-NOV-1998; 98US-0106919.
PR 03-NOV-1998; 98US-0106932.
PR 03-NOV-1998; 98US-0106934.
PR 10-NOV-1998; 98US-0107783.
PR 17-NOV-1998; 98US-0108775.
PR 17-NOV-1998; 98US-0108779.
PR 17-NOV-1998; 98US-0108787.
PR 17-NOV-1998; 98US-0108788.
PR 17-NOV-1998; 98US-0108801.
PR 17-NOV-1998; 98US-0108802.
PR 17-NOV-1998; 98US-0108806.
PR 17-NOV-1998; 98US-0108807.
PR 17-NOV-1998; 98US-0108867.
PR 17-NOV-1998; 98US-0108925.

PR 18-NOV-1998; 98US-0108848.
 PR 18-NOV-1998; 98US-0108849.
 PR 18-NOV-1998; 98US-0108850.
 PR 18-NOV-1998; 98US-0108851.
 PR 18-NOV-1998; 98US-0108852.
 PR 18-NOV-1998; 98US-0108858.
 PR 18-NOV-1998; 98US-0108904.
 (GENTECH) GENENTECH INC.
 PI Baker K, Goddard A, Gurney AL, Smith V, Matanabe CK, Wood WI,
 XX WPI: 2000-237871/20.
 DR P-PSDB; AAY99405.
 XX New mammalian DNA sequences encoding transmembrane, receptor or
 PT secreted PRO polypeptides, useful for screening of potential peptide or
 PT small molecule inhibitors of the relevant receptor/ligand interactions
 XX
 PS Claim 2; Fig 131; 773pp; English.
 XX
 CC AAA37022 to AAA37144 encode the new isolated human transmembrane,
 CC receptor or secreted PRO polypeptides given in AAY99340 to AAY99462. The
 CC transmembrane and receptor PRO proteins can be used for screening of
 CC potential peptide or small molecule inhibitors of the relevant
 CC receptor/ligand interactions. The polypeptides and nucleotide sequences
 CC encoding them have various industrial applications, including uses as
 CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent
 CC PCR primers and hybridisation probes used in the isolation of the PRO
 CC polypeptides from the present invention.
 XX
 SQ Sequence 2848 BP; 607 A; 873 C; 828 G; 540 T; 0 other:
 XX
 Alignment Scores:
 Pred. No.: 2.03e-69 Length: 2848
 Score: 908.50 Matches: 237
 Percent Similarity: 50.51% Conservative: 159
 Best Local Similarity: 30.23% Mismatches: 335
 Query Match: 21.03% Indels: 53
 DB: Gaps: 21
 US-10-025-380-1081 (1-832) x AAA37087 (1-2848)
 QY 62 LeuThrglyglu-----ThraspsanllepheValillegluArgglu---gly 76
 DB 296 CTGTCAGGGGACTCAGGCAAGCACTGAGGCCCATTTGCTATGATCCAGATTCGGC 355
 QY 77 LeuLeuTyrrAaAaArgAlaLeuAAspArggluThraSerThraAsnleuGlnVal 96
 DB 356 TTCCTGCTGCGAGCAGGGCCCTGGACCGAGAGAGCAGAGTACCACTACAGGTC 415
 QY 97 AlaAlaLeuAAspAlaAaAnglylleValleGluGlyProValProleThrilleGluVal 116
 DB 416 ACCCTGGAGATGACGAGATGACATGCTTGGGGGTCCACAGCCTGCTGTCACCTG 475
 QY 117 LysAspIleAaAaAAspAAsnArgProThraPheleuGlnserLysTyrrGluGlySerValArg 136
 DB 476 AAGATGATGAATGACCAAGGTCCTCATTTCTCTCAAGCATCTACAGAGCTCGGTGAC 535
 QY 137 GlnAsnSerArgProGlyLysProPheLeuTyrrValAaAaAlaAaAspLeuAAspPro 156
 DB 536 CGGGGTACAGAGCTGAGATCCCTCTCTCTCTGAGGCTTCAGACGGGATGAGCA 595
 QY 157 AlaThraProAaAnglylleuTyrrGlnleValilleGlnleuProMetilleAaAsn 176
 DB 596 GGCACAGCACTCGGATCTTCATCCATCTCAGACGAGGCTCCAGCCAGCTTCC 655
 QY 177 ValMetTyrrPheGlnlleAaAaAAsnLysThraGlyAlaIleSerleuThraArgGluGlySer 196
 DB 656 CCAGACATGTTCCAGCTGAGCCTCGCTGGGGGCTCGGCCCTCAGCCCAAGGGAGGC 715
 QY 197 GlnGluLeuAaAaProAlaLysAAspProSerTyrrAaAaValilleSerValLysAAspMet 216

DB 716 ACCAGCCTTGACCAAGCCCTGAGAGAGACCTACCACTGTTGGTACAGGTCAAGCATG 775
 QY 217 GlyGlyGlnSerGluAaAAsnSerPheSerAAspThraSerValAspIlelleValThrglu 236
 DB 776 GGTGACCAAGGCC--TCAGGCCACCAAGGCCCATCTGCCACCGTGAAGTCTTCATCATAGAG 832
 QY 237 AsnIleTrpLysAlaProLysProValIleGluMetValIleAaAAsnSerThraProHISPro 256
 DB 833 AGCACTGGGTGTCCTTACAGCTTACCACTGAGAGATCTCAAGGCTATACCCG 892
 QY 257 IleLysIleThrglnValArgTrpAaAAspProGlyAlaGlnTyrrSerleuValAspLys 276
 DB 893 CACCAATGCGCCAGGATGACCTGAGAGTGGGGTGAATGTCACATACCTG----- 943
 QY 277 GlnLysLeuProArgPheProPheSerIleAspGlnGluGlyAspIleTyrrValThrgln 296
 DB 944 GAGAGCCATCCCGGGAGCCCTTGAAGTAAATGCAAGAGGAACCTTACGATGACCA 1003
 QY 297 ProLeuAAspArgGluGlnLysAAspAlaTyrrValPheTyrrAlaValAlaLysAAspGluTyrr 316
 DB 1004 GAGCTGGACAGAGAACCCAGGCTGAGTACCTGCTCAGGTGGGGCTCAGAAATTCCTCAT 1063
 QY 317 GlyLysProLeuSerTyrrProLeuGlnIleHisValLysValLysAspIleAaAAspAAsn 336
 DB 1064 GGGAGAGACTATGCGGCCCTCTGAGCTGACAGTGTGTGTGATGAGATGACACAC 1123
 QY 337 ProProThraCyrrProSerProValThraLysGluValGlnGlnAaAnglylleuGly 356
 DB 1124 GTGCTATCTGCCCCCTCCCTGACCCCAAGTACAGATTCCTTACGTCACGTCCACCAAGGT 1183
 QY 357 AsnSerIleGlyThraLeuThraAlaHisAAspArgAAspGlnLysAaAAsnSerPhe 376
 DB 1184 ACTGAGTGTACTGACTGCTGACAGAGATGACAGATGCCCCGGCTCCCAATTCCTCCAC 1243
 QY 377 LeuAaAAspTyrrArgIleValleGlnIleThraProLysLeuProMetAAspGly---LeuPheleu 395
 DB 1244 GTTGTATATAGCTCTCTGAGCCCTGAGCCCTGAGATGAGGTGATGAGGAGAGACCTTCAG 1303
 QY 396 IleGlnThraTyrrAlaGlyMetleuGlnleuAlaLysleuSerleuLys---GlnAAsp 414
 DB 1304 GTGAGACCCACTTCAGAGCTGACGCTGGGGGTCTCCCATCTCCAGAGGCCACAGAC 1363
 QY 415 ThrProGlnTyrrAaAAsnLeuThraIleGlnValleLysAAsp-----PheLysThraLeu 432
 DB 1364 ATCTGCTTGTGGTGGCTGGCCATGACGTGACGAGGAGCAAGGGTGGCTTCACAGACAGC 1423
 QY 433 CysPheValGlnIleAaAAsnValIleAspIleAaAAspGlnIleProIlePheGluLysSer 452
 DB 1424 TGTGAAGTCCAAAGTCCGACGACACAGATATCAATCAAGATCACCCCTGAGTTCATCACTTCC 1483
 QY 453 AspTyrrGlyAaAAsnLeuThraIleAaAAspThraAsnIleGlySerThraIleleuThraIle 472
 DB 1484 CAGATTGGGCTTAAAGCTTCTCTGAGATGTGAGAGCCGGAGACTCGGTGGCCATGCTA 1543
 QY 473 GlnAlaThraAAspAlaAAsp---GluProPheThrglySerSerLysleu---TyrrHis 490
 DB 1544 AAGACATATGATCTGACCTCGAGGCC-----GCCCTCGGCTCATGGAATTTGGCC 1594
 QY 491 IleleuLysGlyAAspSerGlnGlyArgleuGlyValAspThraAAspProHISThraThra 510
 DB 1595 ATTGAGAGGGGAGACACAGAAAGGAACTTTGGCTGATTTGGAGCA-----GACTCT 1648
 QY 511 GlyTyrrValIleleuLys-----LysProLeuAAspPheGlnThraAlaValleLysAAsn 528
 DB 1649 GGGATGTTGACTGACGACTCTGCAAGAACTTATAGGAGAGCTTCAAGTCATGAG 1708
 QY 529 IleValPheLysAlaGluAaAAspProGluProleuValPheGlyValLysTyrrAaAAspSer 548
 DB 1709 GTGTTGTGGTGGTGGACAGATGTGGGAACTGTG---GGGCAAGGCCCAAGGCGCTTGA 1765
 QY 549 SerPheAlaLysPheThraIleuValleThraAAspAlaAAspGlnIlePheSerGln 568
 DB 1766 GCCACCGCCAGGAGTGTGATGAGAGAGATGTCACACCCCAAGTTGAGACAG 1825

```

QY      569 HisValPheGlnAlaIleValSerGluAspValAlaIleGlyThrIleValIleAsnVal 568
      1826 GAGAGCTACGAGGAGGAGTGTCCCATCATGAGGCCCAAGCCGGCTCTTCTCGTGCATCATC 1885
QY      589 ThrAlaIleAspProGluGlyLeuAspIleSerIleValArgIleAspThrArgGly 608
      1886 CAGCCCTCCAGCCCATCAGCCCAAGCCCTGAGTTCTCTTCTCATGATCATGATGAGAGGC 1945
QY      609 TrpLeuIleAspHisValThrGlyGluIlePheSerValAlaProLeuAsp---Arg 627
      1946 TGCGTCTGCATTGAGAAATCTCCGGAGAGTGCACACCCCAAGTCTTCGACAGGGGCGC 2005
QY      628 GluAlaGlySerProIleArgValGlnValAlaAlaThrGluValGlyIleSerLeu 647
      2006 CAGCTCGGAGACCTTACAGCGTGTGTGAGGCGCCAGATACAGCC----- 2053
QY      648 SerSerValSerGluPheHisLeuIleLeuMetAspValAsnAspAspProAlaGlu 667
      2054 -----CTGACCTCTGCCCCCTGCTCCCA----- 2080
QY      668 AlaIleAspTyrThrGlyLeuPhePheCysHisProLeuSerAlaProGlySerLeuIle 687
      2081 -----TACCTTGACACCCCGCAAGACCATGAGCTTGATGTCGT 2119
QY      688 Phe---GluAlaThrAspAspAspGlnHisLeuPheArgIleProHisPheThrPheSer 706
      2120 AGTGAACCCAGCAAGAGCCCGATCTGCGACGAGGAGCGTCC---TACAGCTTACAC 2176
QY      707 LeuGly---SerGlySerLeuGlnAspAspTyrGluValSerIleAsnGlyThrHis 725
      2177 CTGTGTCACCAACCCACGCGTGAACGCGATTTGGCGCTCAGACTCATGATGTTCCCAT 2236
QY      726 AlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaTyrValValIleuIleArg 745
      2237 GCTTACTCTACCTTGGCCCTGTCATTGGGTGAGGACACGTAACATATCTCCCGTGTG 2296
QY      746 IleAsnAspGlyIleArgProProLeuGluGlyIleValSerLeuProValThrPheCys 765
      2297 GTAGGCACAAATGCCAG-----ATGTGACAGCTCTGTGTGAGATGATCGATGT 2347
QY      766 SerCys---ValGluGlySerCysPheArgProAlaGlyHisGlnThrGlyIleProThr 784
      2348 CGCTGCACCGTGAAGGGGAGTGCATGCCAAGGTGGCGCATGAAAGGCGATGCCACG 2407
QY      785 ValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGlyIleIleuAla 804
      2408 AAGCTGTGGAGTGGGATCTTGTGAGCACTGTGATCAATAGGAATCTTCTCATC 2467
QY      805 ValValPheIleArgIleIleValAspGlyGlyAspAsnValGluSerAlaGlnAla 824
      2468 CTCATTTCACCACTGAGCAATGTCATGAGAAAGAACCCGATCATCAACAGACAGACG 2527
QY      825 SerGluValIle 828
      2528 GTGCCCTGMAAG 2539
DB

```

RESULT 12

```

AA92106
ID AA92106 standard; cDNA, 2848 BP.
XX
AC AA92106;
XX
DT 15-MAY-2001 (first entry)
XX
DE Human PRO1340 cDNA.
XX
KW Human; PRO protein; mapping; ss.
XX
OS Homo sapiens.
XX
PN MO200116318-A2.
XX

```

```

PD      08-MAR-2001.
XX
XX      24-AUG-2000; 2000MO-US23328.
PF
XX
PR      01-SEP-1999; 99MO-US20111.
PR      15-SEP-1999; 99MO-US21090.
PR      07-DEC-1999; 99US-0169495.
PR      11-JAN-2000; 2000US-0170262.
PR      18-FEB-2000; 2000MO-US04341.
PR      18-FEB-2000; 2000MO-US04342.
PR      22-FEB-2000; 2000MO-US04414.
PR      01-MAR-2000; 2000MO-US05601.
PR      03-MAR-2000; 2000US-0187202.
PR      25-APR-2000; 2000US-0193197.
PR      22-MAY-2000; 2000MO-US14042.
PR      05-JUN-2000; 2000US-0209832.
XX
XX      (GETH ) GENENTECH INC.
PA
PI      Eaton DL, Filvaroff E, Gerlisen ME, Goddard A, Godowski PJ;
PI      Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;
XX
XX      WPI, 2001-183260/18.
DR      P-PDB; AAB87574.
XX
XX      Eighty four nucleic acids encoding PRO polypeptides, useful in
PT      molecular biology, including use as hybridization probes, and in
PT      chromosome and gene mapping.
XX
XX      Claim 2; Fig 97; 278pp; English.
XX
XX      The present sequence is the coding sequence for a human PRO polypeptide
CC      (secreted and transmembrane). The PRO protein, and PRO agonists, PRO
CC      antagonists or anti-PRO antibodies are useful for preparation of a
CC      medicament useful in the treatment of a condition which is responsive to
CC      the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO
CC      protein may also be employed as molecular weight markers for protein
CC      electrophoresis. The PRO coding sequence has applications in molecular
CC      biology, including use as hybridisation probes, and in chromosome and
CC      gene mapping.
SQ
XX      Sequence 2848 BP; 607 A; 873 C; 828 G; 540 T; 0 other;
XX

```

Alignment Scores:

```

Pred. No.: 2, 036-69 Length: 2848
Score: 908.50 Matches: 237
Percent Similarity: 50.51% Conservative: 159
Best Local Similarity: 30.23% Mismatches: 335
Query Match: 21.03% Indels: 53
DB: Gaps: 21

```

US-10-025-380-1081 (1-832) x AA92106 (1-2848)

```

QY      62 LeuThrGlyGlu-----ThrAspAsnIlePheValIleGluArgIu---Gly 76
      296 CTGTACGGGAGTCAAGCAAGCACTGAGGCGCCATTGTGATGATCAATCTGCG 355
DB
QY      77 LeuLeuTyrTrpAsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnVal 96
      356 TTCTGCTGTGATCACCGGCGCTGTGACCGAGAGACAGATCAAGTCAAGTCAAGTGC 415
DB
QY      97 AlaAlaLeuAspAlaAsnGlyIleIleValGluGlyProValProIleThrIleGluVal 116
      416 ACCCTGAGATGACAGATGATGATGATCTGTGGGGGTCACAGCCCTGCTTGACGCGT 475
DB
QY      117 LysAspIleAsnAspAsnArgProThrPheLeuGlnSerIleValGluValArg 136
      476 AAGGATGAGAAATGACAGAGTGCCTCATTTCTCAAGCATCTACAGACTGCGTGAAC 535
DB
QY      137 GlnAsnSerArgProGlyIleProPheLeuTyrValAlaAsnAlaThrAspLeuAspPro 156
      536 CGGGATACAGAGGCTGGCATCTCCCTTCTTCTTGAAGGCTTCAAGCCGGATGAGCA 595
DB

```

Oy	157	AlaThrProAngLysGlnLeuThrTyrGlnIleValIleGlnLeuProMetIleAsnAsn	176
Db	596	GGCAGCAGCAACTGGATCTTCATCTCAATCTCGAGCCAGGCTCCAGCCACCCTCC	655
Oy	177	ValMetYrPheGlnIleAsnAsnIleThrGlyAlaIleSerLeuThrArgGlySer	196
Db	656	CCAGACATGTTCCAGCTGAGCCTCCGGCTGGGGGCTCGGCCCTCAGCCCAAGGGAGC	715
Oy	197	GlnGluLeuAsnProAlaLysAsnProSerYrAsnLeuValIleSerValLysAsnMet	216
Db	716	ACCAAGCCTTGACCAAGCCCTGGAGAGAGCACTTACAGCTGTGTGTGATCAGGTCAAGAGCATG	775
Oy	217	GlyGlyGlnSerGluAsnSerPheSerAspThrThrSerValAspIleIleValThrGlu	236
Db	776	GGTGACCAAGGCC--TCAGGGCCACAGGCCACTGCCACCTGGAGAGTCTCATATGAGAG	832
Oy	237	AsnIleTrrpLysAlaProLysProValGluMetValGluAsnSerThrAspProHisPro	256
Db	833	AGCACCCTGGGTGTCCTTACAGCTCAATCCACCTGGAGGAATCTCAAGTCTTATCCCG	892
Oy	257	IleLysIleThrGlnValArgTrrpAsnAspProGlyAlaGlnTyrSerLeuValAspLys	276
Db	893	CACCAACATGCCCCAGGTACACTGAGTGGGGGTATGTGCATCTACCTG-----	943
Oy	277	GluLysLeuProArgPheProPheSerIleAsnGlnGlyAspIleTyrValThrGln	296
Db	944	GAGAGCCATCCCCCGGAGCCCTTTGAATGATGACAGAGGAGAACTCTACGTACACAGA	1003
Oy	297	ProLeuAspArgGluGlnLysAspAlaTyrValPheTyrAlaValAlaLysAspGlnTyr	316
Db	1004	GAGCTGGACAGAGAAGCCCAAGCCTGAGTACCTCGCTCCAGGTGCGGGCTCAAAATTCCTCAT	1063
Oy	317	GlyLysProLeuSerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsn	336
Db	1064	GGCGAGGACATATGGCGGCCCTCTGTGAGCTGACGCGTGTGATGATGAGATATGACAAAC	1123
Oy	337	ProProThrCysProSerProValThrValPheGluValGlnGluAsnGluArgLeuGly	356
Db	1124	GTGCTTATCTACCTCCCTCCCGTGACCCACAGTCACAGATCCCTGAGTCACTCAGTCCACAGGT	1183
Oy	357	AsnSerIleGlyThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPhe	376
Db	1184	ACTGAAGGACTAAGCTCTCAGCAGAGAGATGCATATGCCCGGCTCCCCCAATTCACAC	1243
Oy	377	LeuLeuThrYrArgIleValGlnGlnThrProLysLeuProMetAspGly--LeuPheLeu	395
Db	1244	GTGTGTATCAGCTCTGAGCCCTGAGCCTTGAGTGGGTGAGAGGGAGAGACCTTCCAG	1303
Oy	396	IleGlnThrTyrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLysLys--GluAsp	414
Db	1304	GTGACCCCACTTCAAGGAGGTGACGTGACGCTGGGGGTGCTCCCACTCGAGCAGGCCAGAAC	1363
Oy	415	ThrProGlnTyrAsnLeuThrIleGluValSerAspLysAsp-----PheLysThrLeu	432
Db	1364	ATCTCGCTTCGTGCTGGCCATGAGCACTGGAGCGCAGAGAGGGGTGCTTCAAGAGCAAG	1423
Oy	433	CysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGlnLysSer	452
Db	1424	TGTGAAGTCGAAGTCGACATCAACAATATCATATGTCAGGCCCTGAGTTATCATCTTCC	1483
Oy	453	AspTyrGlyAsnLeuThrLeuAlaGluAspThrAsnIleGlySerThrIleLeuThrIle	472
Db	1484	CAGATGGGCGCATTAAGCTCCCTGAGATGTGAGAGCCCGGAGCTCTGTGGCCATGCTA	1543
Oy	473	GlnAlaThrAspAlaAsp--GluProPheThrGlySerSerLysIleLeu--TyrHis	490
Db	1544	ACAGGCAATGATGCTGACCTCGAGGCC-----GCCCTCCGCTCATGATATTTGGC	1594
Oy	491	IleIleLysGlyAsnSerSerGluLysArgLeuGlyValAspThrAspProHisThrAsnThr	510
Db	1595	ATTGAGAGGAGGACAGAAAGGAGCTTTTGGCTCTGGATTTGGAGCA-----GACTCT	1648

511 G1YrYrAl1le1le1lys-----LyPProLeuAaPheGluThrAla1A1a1SerAa 528
1649 GGGATGATTGACCTCAACACTCTGCAAGAACTCTACGTTATAGAGCAGCTCCAAAGTCAGAG 1708
529 11eValPheValaGluAaProGluProLeuValPheGlyVal1yTyraAla1Ser 548
1709 GTGGAGTGGGCGTGGCAGAGTGGTGGCAACTGGTG---GGGCCAGGCCCAAGGCCCTCGA 1765
549 SerPheAla1yPheThr1leVal1ThraPValAaenGluA1ProGlnPheSerGln 568
1766 GCCACCCGCCAGGTGAGCTGTGCTAGAGAGATGATGACCACCCCAAGTTGGACAG 1825
569 HisValPheGlnAla1yVal1SerGluAaPValAla1leGlyThryA1yVal1G1yAenVal 588
1826 GAGAGCTACAGAGCCAGTGTCCCATAGTGGCCCAAGCCAGGCTTCTTCCTGACCATC 1885
589 ThrAla1yAaPProGluGlyLeuAaP1eSerTySerLeuA1rg1yAaPThrA1rg1y 608
1886 CAGCCCTCCAGCCCATCAGCCCAAGCCCTCAGGTCTCCCTAGTCATGACTCAGAGGGC 1945
609 TrpLeuVal1eAaPHisVal1ThrGlyGlu1lePheSerValAla1aPoleuAaP--Arg 627
1946 TGGCTGTGATGGAAATCTCCGGGAGAGTGACACCCCAAGTCCCTGCGAGGGGCC 2005
628 GluAlaGlySerProTyraArgVal1GlnVal1a1aThrGluVal1Gly1ySerSerLeu 647
2006 CAGCTGGGAGCACCTTACAGAGTCTTGTGGAGGCCCAAGATACAGCC----- 2053
648 SerSerVal1SerGluPheHis1leu1leLeuMetAaPValAaAaPaeProProArgLeu 667
2054 -----CTGACTCTTGCCCTGCTGCCCTCCAA----- 2080
668 Ala1yAaP1yThrGlyLeuPhePheCysHisProLeuSerAla1aPog1ySerLeu1le 687
2081 -----TACCTGTGACACCCGCCCAAGACATGCTGATGTGT 2119
688 Phe---GluAlaThraPaePaeAaGlnHisLeuPheA1rg1yProHisPheThrPheSer 706
2120 AGTGGACCCCAAGAACCCCGATCTGGCCAGTGGGACAGTGGCAGCGATCC--TACAGCTTACC 2176
707 LeuGly---SerGlySerLeuGlnAaAaP1rPog1uVal1SerTy1leAaGlyThryHis 725
2177 CTGTGTCCTCAACCCCAAGTGGACAGGATTTGGGGCTCCAGACTCTCATGTGTTCCAT 2236
726 Ala1rgLeuSerThraArgHis1ThraPaePheGluGluA1rgAla1yVal1a1leu1leA1rg 745
2237 GCTTAACCTCACCTTGGCCCTGATCTGGTGGAGGACATGAACATATAATCCCGTGTGT 2296
746 1leAaAaP1yGlyA1rgProProLeuGluGly1leVal1SerLeuProVal1ThrPheCys 765
2297 GTCCAGCCCAATGCCAG-----ATGTGGCAGTCTCGTGTTCAGAGTATCTGTGT 2347
766 SerCys---Val1GluGlySerCysPheA1rgProAlaGlyHisGlnThrGly1leProThr 784
2348 CGTGTCAACGTGGAGGGGACAGTGCATGCGCAAGTGGGCGGCAAGAGGCAATCCCAAG 2407
785 Val1GlyMetAlaVal1Gly1leLeu1leThrThrThrLeuVal1leGly1le1leLeuAla 804
2408 AAGCGTGGCAGTGGGATCTCTTGTAGGACCTGTGATGAATAGAAATCTTCTCATC 2467
805 ValValPhe1leA1rg1le1y1y1yAaP1y1y1yAaPaeVal1GlySerAlaGlnAla 824
2468 CTCATTTTCCACCACTGGACCATGTCAAGAAAGAAAGACCCGATCAACAGCAGACAGC 2527
825 SerGluVal1y1y 828
2528 GTGCCCTGAG 2539

RESULT 13
AAAF54364
ID AAF54364 standard; DNA; 2648 BP.
XX AAF54364;
AC

XX 02-APR-2001 (first entry)
 XX DNA encoding protein of the invention #66.
 DE Secreted; transmembrane; gene therapy; 88.
 XX Unidentified.
 OS WO200078961-A1.
 XX 28-DEC-2000.
 XX 18-FEB-2000; 2000WO-US04342.
 XX 23-JUN-1999; 99US-0141037.
 XX 20-JUL-1999; 99US-0144758.
 XX 26-JUL-1999; 99US-0145698.
 XX 01-SEP-1999; 99WO-US20111.
 XX 29-OCT-1999; 99US-0162506.
 XX 30-NOV-1999; 99WO-US28313.
 XX 02-DEC-1999; 99WO-US28551.
 XX 16-DEC-1999; 99WO-US30095.
 XX 05-JAN-2000; 2000WO-US00219.
 XX 06-JAN-2000; 2000WO-US00376.
 XX (GETH) GENENTECH INC.
 XX Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Feng S,
 PI Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ,
 PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D,
 PI Watanabe CK, Williams PM, Wood WI;
 DR WPI; 2001-071395/08.
 XX
 PT Secreted and transmembrane proteins and nucleic acids designated PRO,
 PT useful as hybridization probes, in chromosome and gene mapping and gene
 PT therapy -
 XX
 PS Claim 2; Fig 131; 787pp; English.
 XX
 CC The present invention relates to secreted and transmembrane proteins.
 CC These proteins and the DNA encoding them may be used as hybridization
 CC probes, in chromosome and gene mapping and in the generation of
 CC anti-sense RNA and DNA. They may also be used to generate either
 CC transgenic animals or knockout animals which are in turn useful for
 CC development and screening of therapeutically useful reagents.
 CC The nucleic acids may also be used in gene therapy.
 CC
 SQ Sequence 2848 BP; 607 A; 873 C; 828 G; 540 T; 0 other;
 Alignment Scores:
 Pred. No.: 2.03e-69 Length: 2848
 Score: 908.50 Matches: 237
 Percent Similarity: 50.51% Conservative: 159
 Best Local Similarity: 30.23% Mismatches: 335
 Query Match: 21.03% Indels: 53
 DB: 22 Gaps: 21
 US-10-025-380-1081 (1-832) x AAF54364 (1-2848)
 QY 62 TleThrGlyGlu-----ThrAspAsnIlePheValIleGluArgGlu---Gly 76
 DB 296 CTCTCAGGAGACTCAGGCAAGCACTGAGGCGCCATTTCCTATGATCCAGATTCTGCGC 355
 QY 77 LeuLeuTyTyrAsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnVal 96
 DB 356 TTCTCTCTGCTGACCCAGGCGCTGAGCCGAGAGAGAGAGTACCGTACGTCAGGTC 415
 QY 97 AlaAlaLeuAspAlaAsnGlyIleIleValGluGlyProValProIleThrIleGluVal 116
 DB 416 ACCCTGAGATGAGATGAGATGATGCTTGTGTGGGCTCCACGACCTGTGTCAGCTG 475

QY 117 LysAspIleAsnAspAsnArgProThrPheLeuGlnSerLysTyrgluGlySerValArg 136
 DB 476 AAGATGAGAAATATACCGAGTGGCCCATTTCTCTCAAGCCATCTACAGAGTCTGGCTGAGC 535
 QY 137 GlnAsnSerArgProGlyLysProPheLeuTyTyrValAsnAlaThrAspLeuAspArgPro 156
 DB 536 CGGGTACACGAGGCTGGCATCCCTTCTCTTCTTCTGAGGCTTCAGACCGGAGTGAAGCA 595
 QY 157 AlaThrProAsnGlyGlnLeuTyTyrGlnIleValIleGlnLeuProMetIleAsnAsn 176
 DB 596 GGCACAGCCACTCGGATCTTCGATTCACATCTCTGAGCTCAGGCTCCAGCCAGCTTCC 655
 QY 177 ValMetTyPheGlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySer 196
 DB 656 CCAGACATGTTCCAGCTGGAGGCTCGGCTGGGCTCTGGCCCTTCAGCCCAAGGGGAGC 715
 QY 197 GlnGluLeuAsnProAlaLysAsnProSerTyTyrAsnLeuValIleSerValLysAspMet 216
 DB 716 ACCAGCTTGACCAACCCCTGGAGAGAGACTTACAGCTGTGTTACAGTCAAGACATG 775
 QY 217 GlyGlyGlnSerGluAsnSerPheSerAspThrThrSerValAspIleIleValThrGlu 236
 DB 776 GGTGACCAAGGCT--TCAGGCCACAGGCCACTGCCACGTTGAGATCTTCATCATAGAG 832
 QY 237 AsnIleTyPheValAlaProLysProValGluMetValGluAsnSerThrAspProHisPro 256
 DB 833 AGCACCTGGGTGTCCTTACAGCTTATCCACCTGGCAGAGATCTCAAGCTCATATCCG 892
 QY 257 IleLysIleThrGlnValArgTrpAsnAspProGlyAlaGlnTySerLeuValAspLys 276
 DB 893 CACCAATGCGCCAGGTATCACTGAGTGGGGGATGATGCATATCACTG-----943
 QY 277 GluLysLeuProArgPheProPheSerIleAspGlnGlyAspIleTyTyrValThrGln 296
 DB 944 GAGAGCATTCCTCCGGACCTTTGAGTGAATGACAGAGGAAACCTTACGTGACCAAC 1003
 QY 297 ProLeuAspArgGluGluLysAspAlaTyTyrValPheTyTyrAlaValAlaLysAspGluTyr 316
 DB 1004 GAGCTGACAGAGAAACCCAGGCTGAGTACCTCTCCAGGTGGGCTCAGATTTCCTCAT 1063
 QY 317 GlyLysProLeuSerTyTyrProLeuGlnIleHisValIleValLysAspIleAsnAspAsn 336
 DB 1064 GCGGAGACTATGACGCGCCCTGAGCTGACGTGCTGATGATGAGATGACCAAC 1123
 QY 337 ProProThrCysProSerProValThrValPheGluValGlnGluAsnGluArgLeuGly 356
 DB 1124 GTGCTATCTGCTCCCTCCGTCGACCCACAGTCAGATCCGTCAGTCCAGTCCACAGGT 1183
 QY 357 AsnSerIleGlyThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPhe 376
 DB 1184 ACTGAAGTACATGACTGTCTCAGCAGAGATGACGATGCCCGGCTCCCAATTCTCCAC 1243
 QY 377 LeuAsnTyTyrArgIleValGluGlnThrProLysLeuProMetAspGly--LeuPheLeu 395
 DB 1244 GTTGTATACAGCTCTGAGCCCTGAGCTGAGAGATGGGATGAGGAGAGAGCTTCCAG 1303
 QY 396 IleGlnThrTyTyrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLys--GlnAsp 414
 DB 1304 GTGAGCCCACTTCAGAGAGTGTGACGTGGGGGTCTCCACTCGACAGGCGCAGAAC 1363
 QY 415 ThrProGlnTyTyrAsnLeuThrIleGluValSerAspLysAsp----PheLysThrLeu 432
 DB 1364 ATCTGCTCTGTGCTGTGCTGAGCTGACCTGGCAGGCGCAGAGGCTGCTTCACAGCAAG 1423
 QY 433 CysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLysSer 452
 DB 1424 TGTGAAGTGAAGTGCAGCTCAGATATCATATGATCAAGCCCTGAGTTCACTCATCTCC 1483
 QY 453 AspTyTyrGlyAsnLeuThrLeuAlaGluAspThrAsnIleGlySerThrIleLeuThrIle 472
 DB 1484 CAGATTGGGCTTATTAAGCTCTCCCTGAGATGTGAGCCCGGAGACTTGTGTGCCATGCT 1543
 QY 473 GlnAlaThrAspAlaAsp---GluProPheThrGlySerSerLysIleLeu---TyrHis 490

```

Db      1544  AACGACATGATGCTGACCTGAGCC-----GCTCCGCCCATGAGATTGCC 1594
Qy      491  TLeileylglaSerglglYArgleuGlYAlaPthraSPRoH1sThraSPthR 510
Db      1595  ATTGAGAGGGAGACAGAGAAGGAGCTTTGGCTGATGGAGCCA-----GACTCT 1648
Qy      511  GLYrYrValIlelelyS-----LyeProleuSphRgIuThraIAlaValaSerAsn 528
Db      1649  GGGCATGTGAGTACAGACTCTGCAAGAACTCATATGAGGAGCATCCAGTCATGAG 1708
Qy      529  TLevalPheylalaGluSAnPProgluPProleuValPheglYVallySTYrAsnAlaSer 548
Db      1709  GTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1765
Qy      549  SerPheAlaYSPheThrLeuIleValAlThraSPValaengluAlaPProgluPheSerGln 568
Db      1766  GCCACCCGCCACGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1825
Qy      569  HisValPheglAlaYSPheValSerGluSAspValAlaIleglYThrlYSPValGlyAsnVal 588
Db      1826  GAGAGCTACGAGGCGCTGCTCCATGATGCTCCAGCCGCGCTCTTCTGCTGACCATC 1885
Qy      589  ThrAlaYSPProgluGlYLeuAspIleSerTySerleuArgGlyAspThraSPthRArgGly 608
Db      1886  CAGCCCTCCGACCCCATCAGCCGAAACCTCAGTTCCTCCATGATGATGATGATGATG 1945
Qy      609  TrpleuYSPleAspH1sValThrGlyGluIlePheSerValAlaPProleuAsp--Arg 627
Db      1946  TGGCTGTGATGAGAAATTCCTCCGGGAGGTGACACCCGCCCATGCTGTGACAGGCGCC 2005
Qy      628  GluAlaGlySerProTyTrArgValAlaValAlaThraGluValGlyGlySerleu 647
Db      2006  CAGCTGTGAGACCTTACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2053
Qy      648  SerSerValSergluPheH1sLeuIleleuMetAspValaSPAsnProProArgleu 667
Db      2054  -----CTGACTTTCCTCCCTGTGCTCCCA----- 2080
Qy      668  AlAlaYSPThrThrGlyLeuPhePheCyH1sProleuSerAlaPProGlySerleuIle 687
Db      2081  -----TACCTGTGACACCCGCCCAAGACATGAGCTTGTGTGTGTGTGTGTGTGT 2119
Qy      688  Phe---GluAlaThraSPAspGlnH1sleuPheArgGlyProH1sPheThraSPheSer 706
Db      2120  AGTGAACCCAGCAAGACCCGATCTGCGCATGAGGAGCGGTCCC---TACAGCTTCAAC 2176
Qy      707  LeuGly---SerglySerleuGlnAsnAspTrpGluValSerlySleAsnGlyThraH1s 725
Db      2177  CTGTGTCCCAACCCCAAGGTGCAACGGGATTTGGCGCTCCAGACTCTCAATGATGTTCCAT 2236
Qy      726  AlAlaYSPSerThraSPH1sThraSPhegluGluAlaYrValValleuIleArg 745
Db      2237  GCCTACCTCACTGCGCTGCTGATGAGGACCATGAAACATATATCCCGTGTGTGTGTGT 2296
Qy      746  TLeasnaSPglYglYArgProProleuGlYleValSerleuPProvalThraPheCyS 765
Db      2297  GTGCGCAACATGCCAG-----ATGCGCAGAGCTCTGCTTCAAGATGATGTGTGT 2347
Qy      766  SerCyS---ValgluGlySerCySPheArgProAlaGlyH1sGlnThraGlyIleProThr 784
Db      2348  CGCTGCAACGTGAGAGGGGAGCATGTCGCAAGGTGGGCCCAAGAGGAGCATGCCACG 2407
Qy      785  ValgluMetAlaValglYleleuIleuThrThraSPleuValIleGlyIleIleuAla 804
Db      2408  AAGCTGTGCGAGGTGGGATCTTGTAGGACCTGTGTAGCAATAAGATTTCTCTATC 2467
Qy      805  ValAlaPheIleArgIleleYleYSPheglYSPheValaSPAsnValGluSerAlaGlnAla 824
Db      2468  CTCATTTTCAACCACTGAGCATGTCAGAAAGAGACCCGATCAACAGCAGACAGC 2527
Qy      825  SerGluVallyS 828

```

```

Db      2528  GTGCCCTGAG 2539
RESULT 14
ABK33629
ID      ABK33629 standard; cDNA, 2848 BP.
XX      ABK33629;
AC      ABK33629;
XX      08-MAY-2002 (first entry)
DT      cDNA encoding human PRO protein, Seq ID No 187.
DE      cDNA encoding human PRO protein, Seq ID No 187.
XX      Human; secreted protein; PRO; tumour; lung cancer; colon cancer;
KW      breast cancer; prostate tumour; rectal tumour; liver tumour;
KW      pericyte cell proliferation; chondrocyte cell proliferation;
KW      tumour necrosis factor-alpha; gene; ss.
XX      Homo sapiens.
XX      WO200208288-A2.
XX      31-JAN-2002.
XX      29-JUN-2001; 2001WO-US21066.
XX      20-JUL-2000; 2000US-219556P.
XX      25-JUL-2000; 2000US-220585P.
XX      25-JUL-2000; 2000US-220605P.
XX      25-JUL-2000; 2000US-220607P.
XX      25-JUL-2000; 2000US-220624P.
XX      25-JUL-2000; 2000US-220638P.
XX      25-JUL-2000; 2000US-220644P.
XX      25-JUL-2000; 2000US-220666P.
XX      25-JUL-2000; 2000US-220683P.
XX      25-JUL-2000; 2000US-220710P.
XX      25-JUL-2000; 2000US-220711P.
XX      23-AUG-2000; 2000WO-US23522.
XX      24-AUG-2000; 2000WO-US23528.
XX      15-SEP-2000; 2000US-000000P.
XX      10-NOV-2000; 2000WO-US30873.
XX      28-NOV-2000; 2000US-253646P.
XX      01-DEC-2000; 2000WO-US32678.
XX      20-DEC-2000; 2000US-0747259.
XX      20-DEC-2000; 2000WO-US34956.
XX      28-FEB-2001; 2001WO-US06520.
XX      10-MAY-2001; 2001US-0854280.
XX      25-MAY-2001; 2001WO-US17092.
XX      (GENT ) GENENTECH INC.
XX      Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ,
PI      Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX      WPI; 2002-172001/22.
XX      P-PSDB; AUB3685.
XX      One hundred and twenty two nucleic acids encoding PRO polypeptides,
PT      useful for treating a PRO related disorder and for diagnosing tumours
PT      such as lung cancer, colon cancer, breast tumour, prostate tumour, rectal
PT      tumour or liver tumour -
XX      Claim 2; Figure 187; 359pp; English.
XX      The invention relates to one hundred and twenty two nucleic acids
XX      encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides
XX      encode human secreted proteins. The PRO nucleic acids, polypeptides,
XX      agonists and antagonists are useful for treating a PRO related disorder.
XX      The PRO polypeptides are useful for diagnosing tumours, especially lung
XX      cancer, colon cancer, breast tumour, prostate tumour, rectal tumour or
XX      liver tumour. The PRO polypeptides are useful for stimulating the
XX      proliferation of, or gene expression, in pericyte cells, for stimulating
XX      the proliferation or differentiation of chondrocyte cells, for
XX      stimulating the release of tumour necrosis factor-alpha from human blood,
XX      for stimulating or inhibiting the proliferation of normal human dermal

```

CC fibroblast cells. The PRO polypeptide may also be used as molecular weight markers and for tissue typing. The PRO nucleic acids have applications in molecular biology, including use as hybridisation probes, CC and in chromosome and gene mapping. ABK3536-ABK3657 represent human CC PRO protein coding sequences of the invention.

XX Sequence 2848 BP; 607 A; 873 C; 828 G; 540 T; 0 other;

Alignment Scores:

Pred. No.:	2,036-69	Length:	2848
Score:	906.50	Matches:	237
Percent Similarity:	50.51%	Conservative:	159
Best Local Similarity:	30.23%	Mismatches:	335
Query Match:	21.03%	Indels:	53
DB:	24	Gaps:	21

US-10-025-380-1081 (1-832) x ABK33629 (1-2848)

```

QY 62 LeuThrGlyLyu-----ThrAspAsnIlePheValIleGluArgGlu--Gly 76
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
296 CTGTCAAGGGGACTCAGCAAGGCAAGCACTGAGGGCCATTGCTATGATCAGCTATCTGCG 355
QY 77 LeuLeuTyrrTrsAsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnVal 96
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
356 TTCCTGTGTGTGACCAAGGGCCCTGTGACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 415
QY 97 AlaAlaLeuAspAlaAsnGlyIleIleValGluGlyProValProIleThrIleGluVal 116
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
416 ACCCTGAGATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 475
QY 117 LysAspIleAsnAspAsnArgProThrPheLeuGlnSerLysTyrrGluGlySerValArg 136
DB 476 AAGAGATGAGATGATGACCAAGTGGCCCATTTCTCTCAAGCATCTACAGAGCTCGGCTGAG 535
QY 137 GluAsnSerArgProGlyLysProPheLeuTyrrValAsnAlaThrAspLeuAspAspPro 156
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
536 CGGAGGTACCAAGGCTGTGATCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 595
QY 157 AlaThrProAsnGlyGlnLeuTyrrGlnIleValIleGlnLeuProMetIleAsnAsn 176
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
596 GGCACAGCCAACTCGCATCTTCATCTCAGATCTGAGCCAGGCTCCAGCCAGCCCTTC 655
QY 177 ValMetTyrrPheGlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySer 196
DB 656 CCAGACATGTTCCAGCTGAGAGCTCGGCTGGGGGCTGTGGCTCCAGCCCAAGGGAGAG 715
QY 197 GluGluLeuAsnProAlaLysAsnProSerTyrrAsnLeuValIleSerValLysAspMet 216
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
716 ACCAGCCTTGACCAAGCCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 775
QY 217 GlyGlyGlnSerGlnAsnSerPheSerAspThrThrSerValAspIleIleValIleThrGlu 236
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
776 GTGACCAAGGCC---TCAGGCCACACAGGCCACCTGCAAGCTTCATCATATAGAG 832
QY 237 AsnIleTyrrLysAlaProLysProValGluMetValGluAsnSerThrAspProHisPro 256
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
833 AGCAGCTGGGTGTCTCCAGAGCCTATCCACCTGCGAGAGAGAGAGAGAGAGAGAGAGAG 892
QY 257 IleLeuIleThrGlnValArgTrpAsnAspProGlyValIleGlnTyrrSerLeuValAspLys 276
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
893 CACACATGGCCCGAGGTACCTGAGAGGGGGTGAATGTGACATATACCTG----- 943
QY 277 GluLysLeuProArgTrpPheProPheSerIleAspGlnGluGlyAspIleTyrrValIleThrGln 296
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
944 GAGAGCATATCCCCGGGAGCCCTTTGAAGTAAATGCAAGAGGAGAAACCTTACGTGACCA 1003
QY 297 ProLeuAspArgGlnGluLysAspAlaTyrrValPheTyrrAlaValAlaLysAspGluTyrr 316
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1004 GAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1063
QY 317 GlyLysProLeuSerTyrrProLeuGlnIleHisValLysValLysAspIleAsnAspAsn 336
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1064 GGGAGAGACTATAGGGCCCTCTGTGAGCTGACAGTGTGTGATGATGAGATGACAAAC 1123

```

```

QY 337 ProProThrCysProSerProValThrValPheGluValGlnGluAsnGluArgLeuGly 356
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1124 GTGCTATATCGCCCTCCCGGTGAGCCCAAGCAGTACAGATCTCTAGCTCATCCACAGAGT 1183
QY 357 AsnSerIleGlyThrLeuThrAlaHisAspAspAspGlnGluAsnThrAlaAsnSerPhe 376
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1184 ACTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1243
QY 377 LeuAsnTyrrArgIleValGluGlnThrProLysLeuProMetAspGly---LeuPheLeu 395
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1244 GTTGTATATACGCTCTGAGCCCTGAGCCTGAGAGATGGGGTGAAGGGAGAGAGCTTTCAG 1303
QY 396 IleGlnThrTyrrAlaGlyMetLeuGlnLeuValAlaGlnSerLeuLys---GlnAsp 414
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1304 GTGAGACCCACTTCAGAGAGTGAAGCTGGGGGTCTCCCATCTCCAGAGAGCCAGAAC 1363
QY 415 ThrProGlnTyrrAsnLeuThrIleGluValSerAspLysAsp-----PheLysThrLeu 432
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1364 ATCTGCTTCTGTGTGCTGGCATGAGACTGGCAGAGAGGGGTGAGCTTCAACAGACAG 1423
QY 433 CysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLysSer 452
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1424 TGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1483
QY 453 AspTyrrGlyAsnLeuThrLeuAlaGluAspThrAsnIleGlySerThrIleLeuThrIle 472
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1484 CAGATTGGGCTTAAAGCTCTCCGAGATGTGAGACCCGGAGACTGTGGTGGCAGCTGA 1543
QY 473 GlnAlaThrAspAlaAsp---GluProPheThrGlySerSerLysIleLeu---TyrrHis 490
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1544 ACAGCATTGATGCTGACCTCGAGCCC-----GCCCTCGGCTCATGATTTGGCC 1594
QY 491 IleIleLysGlyAspSerGlnGlyArgLeuGlyValAspThrAspProHisThrAsnThr 510
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1595 ATTGAAGGGAGAACACAGAAAGGACTTTGGCTGATTTGGAGCA-----GACTCT 1648
QY 511 GlyTyrrValIleLys-----LysProLeuAspPheGluThrAlaAlaValSerAsn 528
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1649 GGCAGATTAGACTGACACTGCAAGAACTCAGTATAGGAGAGAGCTCCAAAGTCAATGAG 1708
QY 529 IleValPheLysAlaGluAsnProGluProLeuValPheGlyValLysTyrrAsnAlaSer 548
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1709 GTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1765
QY 549 SerPheAlaLysPheThrLeuIleValIleThrAspValAsnGluAlaProGlnPheSerGln 568
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1766 GCCACCGCCACGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1825
QY 569 HisValPheGlnAlaLysValSerGluAspValAlaIleGlyThrLysValGluAsnVal 588
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1826 GAGAGCTAGAGAGCCAGTGTCTCCATCAAGTCCCAAGCCGAGGCTCTTCTGTGACCATTC 1885
QY 589 ThrAlaLysAspProGlnGlyLeuAspIleSerTyrrSerLeuArgGlyAspThrArgGly 608
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1886 CAGCCTCGAGCCCATCAAGCCGAGAACCTCAGGTTCTCTGTATCATAGTAGAGGG 1945
QY 609 TrpLeuLysIleAspHisValThrGlyGluIlePheSerValAlaProLeuAsp---Arg 627
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1946 TGGCTGTGATGAGAAATTCCTCGGGAGGTGCACACCCAGTCCCTGCAAGGGCGCC 2005
QY 628 GluAlaGlySerProTyrrArgValGlnValValAlaIleThrGluValGlyLysSerLeu 647
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
2006 CAGCTCGGAGACCTACAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2053
QY 648 SerSerValSerGluPheHisLeuIleLeuMetAspValAsnAspAsnProArgLeu 667
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
2054 -----CTGACTCTTGTGCCCCGTGTGCTCCCAA----- 2080
QY 668 AlaLysAspTyrrThrGlyLeuPhePheCysHisProLeuSerAlaProGlySerLeuIle 687
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
2081 -----TACCTGTGACACCCCGCCCAAGACCATGAGCTGTGATGCTG 2119

```

QY 688 Phe---GluAlaThrAspAspArgGlnHisLeuPheArgGlyProHisPheThrPheSer 706
 DB 2120 AGTGGACCCGCAAGAGCCCCGATCTGCGCAAGTGGGACGCTCC--TACAGTTTACC 2176
 QY 707 LeuGly---SerGlySerLeuGlnAsnAspTrpGluValSerLysIleAsnGlyThrHis 725
 DB 2177 CTGTGTCACCAACCCACGAGTGCACGGGATTTGGCGCTCCAGACTCTCATGTGTTCCAT 2236
 QY 726 AlaArgLeuSerThrArgHisIleThrAspPheGluGluArgAlaValIleValIleLeuArg 745
 DB 2237 GCGTACCTGACCTTGGCCCTGTCATTTGGTGGAGCCAGTAAACATTAATCCCGTGTG 2296
 QY 746 ILAsnAspGlyArgPProPLeuGluGlyIleValSerLeuProValThrPheCys 765
 DB 2297 GTGAGCCACATGCCCCAG-----ATGTCGACGCTCTGCTGTCAGATGATGTGT 2347
 QY 766 SerCys---ValGluGlySerCysPheArgProAlaGlyHisGlnThrGlyIleProThr 784
 DB 2348 CGGTGCAACGTGGAGGGGAGTGCATGCGCAAGTGGGCGCATGAAAGGCGATCCACG 2407
 QY 785 ValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGlyIleIleLeuAla 804
 DB 2408 AAGCTGTGGCAGTGGGATCTTGTAGGCGACCTGTGTGCAATAGCAATCTTCTCATC 2467
 QY 805 ValValPheIleArgIleLysLysAspLysGlyLysAspAsnValGluSerAlaGlnAla 824
 DB 2468 CTCATTTCACCCCTGACCTGACCATGTCAAGAAAGAACCCGATCAACGACGACAGC 2527
 QY 825 SerGluValLys 828
 DB 2528 GTGCCCTGAAG 2539
 RESULT 15
 AAS15602
 ID AAS15602 standard; CDNA; 2938 BP.
 AC AAS15602;
 DT 18-JUN-2002 (first entry)
 DE Human CDNA encoding cadherin-2 (CDHN-2).
 KW ss: cadherin; CDHN-2; nootropic; neuroprotective; cardiant; human;
 KW antiinflammatory; gastric; immunostimulant; cyrostatic; gene-therapy;
 KW Alzheimer's disease; Parkinson's disease; multiple sclerosis; cancer;
 KW atherosclerosis; angiogenesis; musculoskeletal; ataxia; myotonia;
 KW gastrointestinal; gastritis; insulin dependent diabetes mellitus;
 KW Crohn's disease; inflammatory; asthma; rheumatoid arthritis; lupus.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX 5'UTR 1..161
 XX /tag= a
 XX sig_peptide 162..224
 XX /tag= b
 XX /note= "signal peptide"
 XX CDS 162..2654
 XX /tag= c
 XX /product= "CDHN-2"
 XX /note= "This sequence is specifically claimed in
 claim 1"
 XX mat_peptide 225..2651
 XX /tag= d
 XX /note= "Mature CDHN-2"
 XX 3'UTR 2655..2938
 XX /tag= e
 XX WO200179293-A2.
 XX 25-OCT-2001.
 XX 18-APR-2001; 2001WO-US12687.

XX
 PR 18-APR-2000; 2000US-198466P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Curtis RAJ;
 DR WPI; 2001-611722/70.
 DR P-FSDB; AAU09959.
 XX
 PT Isolated cadherin polypeptide useful for the treatment of central
 PT nervous system, cardiovascular, musculoskeletal, gastrointestinal,
 PT inflammatory, immune system and cell proliferation disorders -
 XX
 PS Claim 1; Figure 7; 147pp; English.
 XX
 CC This sequence represents the CDHN-2 gene of the invention, this
 CC invention comprises the nucleotide and protein sequences of two
 CC members of the cadherin family of molecules CDHN-1 and CDHN-2.
 CC Also included in the invention are a vector containing the CDHN
 CC gene, a host cell transfected with the vector and a method for
 CC producing the CDHN proteins by culturing the host cell line. Also
 CC included is an antibody which selectively binds to the protein
 CC The CDHN-1 protein may have nootropic, neuroprotective, cardiant,
 CC antiinflammatory, gastric, immunostimulant and cytostatic activity
 CC and can be used as a cadherin-modulator using antisense-therapy or
 CC gene-therapy. These proteins are useful for identifying compounds
 CC which bind or modulate CDHN-1. The proteins, nucleic acids and
 CC antibodies are useful for treating a subject with a disorder
 CC characterised by aberrant or unwanted cadherin protein or nucleic
 CC acid activity. These disorders include central nervous system
 CC (e.g. Alzheimer's disease, Parkinson's disease, multiple sclerosis),
 CC cardiovascular (e.g. atherosclerosis, angiogenesis), musculoskeletal
 CC (e.g. ataxia, myotonia) gastrointestinal (e.g. gastritis, insulin
 CC dependent diabetes mellitus, Crohn's disease) (e.g. inflammatory and
 CC immune system (e.g. asthma, rheumatoid arthritis, lupus) or cell
 CC proliferation disorders (e.g. cancers, leukaemia).
 XX
 SQ Sequence 2938, BP; 723 A; 847 C; 766 G; 602 T; 0 other;
 Alignment Scores:
 Pred. No.: 8,896-67 Length: 2938
 Score: 878.50 Matches: 248
 Percent Similarity: 47.72% Conservative: 140
 Best Local Similarity: 30.50% Mismatches: 358
 Query Match: 20.33% Indels: 67
 DB: Gaps: 22
 US-10-025-380-1081 (1-832) x AAS15602 (1-2938)
 QY 62 LeuThrGlyGluThrAsp-----AsnIlePheValIleGluArgGlu---Gly 76
 DB 336 CTATCAGGAGACTCAACACGCGAGATCAAAACCTTTGCTGTGGACACAGCTGTGGC 335
 QY 77 LeuLeuTyrrTyraAsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnVal 96
 DB 396 TTTCTAGTGGCGACAGGACCTCGAACCGGAGAGAGAACAATATCAACATACAGGTC 455
 QY 97 AlaAlaLeuAspAlaAsnGlyIleIleValGluGlyProValProIleThrIleGluVal 116
 DB 456 ACCTTGAGTCTGAGATGACGATCTTCTGTGGGTCACAGCTTGACTGTGCAATGTG 515
 QY 117 LysAspIleAsnAspAsnArgProThrPheLeuGlnSerLysTyrrGluGlySerValArg 136
 DB 516 AAGATGAGATATACAGTACCCCAATTTCCAGGCACTTACAGAGCTCAGCTGAGC 575
 QY 137 GlnAsnSerArgProGlyLysProPheLeuTyrrValAsnAlaThrAspLeuAspPro 156
 DB 576 CAGGACACGAGGCTGGGGTCCCTTCTCTTCTGAGGCTCTGATGGGAGTACCA 635
 QY 157 AlaThrProAsnGlyLeuTyrrTyrrGlnIleValIleGlnLeuProNecIleAsnAsn 176
 DB 636 GGCACAGCTAACTCCGACCTTCCATTCATGTGAGCGACAGTCCCACTCAGCCTTTA 695

QY 177 ValMetYrPheGlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGlySer 196
 DB 696 CCAGACATGTTCCAGCTGAGACCCCTCACCTAGGGGCTCTGGCTTTATCCACAGTGAAGC 755
 QY 197 GlnGluLeuAsnProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMet 216
 DB 756 ACCAGCTTACCATCCATGGCTTGAAGACCTTACCATGGCTTGAAGACCTTGAAGACCTG 815
 QY 217 GlyGlyValIleSerGluAsnSerPheSerAspThrThrSerValAspIleIleValThrGlu 236
 DB 816 GGGACACAGCTTCCAGGCCACACAGCTATTGCACT--GTAGAGATCTCCATATGTAAG 872
 QY 237 AsnIleThrLysAlaProLysProValGluMetValGluAsnSerThrAspProIlePro 256
 DB 873 AACAGCTGGGACCCCTGAGAGCTTTCACCTGGCAGAGATCCAAAGTTGGTACCA 932
 QY 257 IleLysIleThrGlnValArgTyrPheAsnProGlyValIleGlnTyrSerLeuValAspLys 276
 DB 933 CACAGCATTCGCCAGGTGACCTGAGTGAAGAGACCTGCATCACCTGAGCTG----- 983
 QY 277 GluLysLeuProArgPheProPheSerIleAspGlnGluLysAspIleTyrValThrGln 296
 DB 984 GAGAGCCAGCTCCAGGACCTTCGATGTGATACAGAGGGAGATGCTCCATGTTCATCCATG 1043
 QY 297 ProLeuAspArgGlnGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyr 316
 DB 1044 GAGCTGGACCGGAGGCGGAGCTGAGTACACAGCTCCAACTCCGAGCTCGAATTCCTCAT 1103
 QY 317 GlyLysProLeuSerTyrProLeuGluIleIleValLysValLysAspIleAsnAspAsn 336
 DB 1104 GGTGAGACTACACAGAACCCCTGAGTTGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1163
 QY 337 ProProThrCysProSerProValThrValPheGlnValGlnGluAsnGluLysGly 356
 DB 1164 GCACCTGTCTGCTCCCAATGACCCAACTGACAACTGACAACTGACAACTGACAACTGACAA 1223
 QY 357 AsnSerIleGlyThrLeuThrAlaIleAspArgAspGluLysAsnThrAlaAsnSerPhe 376
 DB 1224 ACTGAATAGCCAGGCTCTGACGAGAGATTTGGATGCTCCCTGGGTCCACCATTCCTCC 1283
 QY 377 LeuAsnTyrArgIleValGluGlnThrProLysLeuProMetAspGly--LeuPheLeu 395
 DB 1284 ATTGTATATCATGTTGTTGCTGAGCCCTGAGCTGAGAGGGGCTGAAACAAAGCTTCGAG 1343
 QY 396 IleGlnThrTyrIleGlyMetLeuGlnLeu-----AlaLysGlnSer 409
 DB 1344 TTGATCCGACCTCAGGACAGTGTACACTGGAACTGCCCACTCATGCTGGCCAGAGT 1403
 QY 410 LeuLysLeuGlnAspThrProGlnTyrAsnLeuThrIleGluVal-----SerAspLys 427
 DB 1404 ATCTGCTCTCAG-----GTCTGCTGCTTGACTTACCTAGCAGGATCAGAGAGT 1448
 QY 428 AspPheLysThrLeuCysPheValGlnIleAsnValIleAspIleAsnAspGlnIlePro 447
 DB 1449 GGCCTCAGACGACATGTGATGAGCAGTCACTGATGACAGCTCAACAAACATGCCCCCT 1508
 QY 448 IlePheGlnLysSerAspTyrGlyAsnLeuThrLeuAlaGluAspThrAsnIleGlySer 467
 DB 1509 GAGTTATCATATCCAGATGGGCTGTAATCTTCTGAGAGTGTAAACCTGGGGCT 1568
 QY 468 ThrIleLeuThrIleGlnAlaThrAspAlaAsp--GluProPheThrGlySerLys 486
 DB 1569 CTGGTGGCAACTCATGCGCACTGATGCTGACCTTGAACT-----GCTTCCCGC 1619
 QY 487 IleLeu--TyrHisIleIleLysGlyAspSerGlnGlyArgLeuGlyValAspThrAsp 505
 DB 1620 CTATGACATTTGCTGATGAAGAAGAGACCCAGAGGATCTTTGACCTTGCTGGAG 1679
 QY 506 ProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAspPheGlnThrAlaAla 525
 DB 1680 CCAAGCTCCGACCATGTCCAGCTCAGATCCGGAAAGAACTCATATGAGGAGGACTCTCT 1739

QY 526 ValSerAsnIleValPheLysAlaGluAsnProGluProLeuValPheGlyValLysTyr 545
 DB 1740 GATCACAAGGTGTGTGGTGTGATGATTAACATAGAAACATGTGTG--GGCCAGGGCCCA 1796
 QY 546 AsnAlaSerSerPheAlaLysPheThrLeuIleValIleThrAspValAsnGluAlaProGln 565
 DB 1797 GGCCCTGACGCCACACACAGTCACTATCATGTGAGAGGGGTGTGTGTCTCCCTCAAG 1856
 QY 566 PheSerGlnHisValPheGlnAlaLysValSerGluAspValAlaIleGlyThrLysVal 585
 DB 1857 TTGACACAGAGACTATGAGACACACATCCAGTACAGACCCCACTGGCTCTCTCTG 1916
 QY 586 GlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTyrSerLeuAspGlyAsp 605
 DB 1917 CTGACCATCAGCCCTCAGACCCCATAGAGAGAACCTCAGTTCCTCGTGCATATGAC 1976
 QY 606 ThrArgGlyTyrPheLysIleAspHisValThrGlyGluIlePheSerValAlaProLeu 625
 DB 1977 TCAGAGGGCTGGCTGTATCAAGAGGTGTCTGGGAGGTACACAGCCAGTCCCTG 2036
 QY 626 Asp--ArgGluAlaGlySerProTyrArgValGlnValAla-----ThrGluVal 642
 DB 2037 CAGGCTGCCAGCTGAGACACATACATACAGTCTTGTGAGGCTCCAAAGACAGATTAAG 2096
 QY 643 GlyLysSerSerLeuSerSerValSerGluPheIleLeu----- 656
 DB 2097 CCAAGACTAGACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2156
 QY 657 LeuMetAspValAsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePhe 676
 DB 2157 GCATTGACTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2192
 QY 677 CysHisProLeuSerLysProGlySerLeuIlePheGluAlaThrAspAspAspGlnHis 696
 DB 2193 TGTACACCCCGCCAGACATACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2252
 QY 697 LeuPheArgGlyProHisPheThrPheSerLeuGly--SerGlySerLeuGlnAsnAsp 715
 DB 2253 GCCAAGAGAAATGTGCTCCCAAGCTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2312
 QY 716 ThrGluValSerLysIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPhe 735
 DB 2313 TGGCGCTCAGCCCTTCAACATTCACAGCTTACCTGACCTTGGCATTTGATGGGTA 2372
 QY 736 GluGluArgAlaTyrValValLeuIleArgIle--AsnAspGlyLysArgProProLeu 754
 DB 2373 GAGCTGTGAAATACATGTGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2432
 QY 755 GluGlyIleValSerLeuProValThrPheCysSerCys--ValGluGlySerCysPhe 773
 DB 2433 CAG-----GTCAAAGT 2480
 QY 774 ArgProAlaGlyHisGlnThrGlyIleProThrValGlyMetAlaValGlyIleLeuLeu 793
 DB 2481 CCCAAGTGGTGTGATGAAGGAATGCCCCAGAACTGTCAAGCGGTGTGTGTGTGTGTGTGT 2540
 QY 794 ThrThrLeuLeuValIleGlyIleIleLeuAlaValAlaPhe-----Ile 808
 DB 2541 GGCACCTGTGACAGCATGAGTTCATTTCTCATTTCTGTTCACGACCTGTGGCCCTGGA 2600
 QY 809 ArgIleLysLysAspLysGlyLysAspAsnVal----- 819
 DB 2601 AGGAAGACCTGTGATCAGCAGACAGACAGCCTGCTGTAAGCAGCGGTGTGAATGATC 2660
 QY 820 GluSerAlaGlnAlaSerGluValLysProLeuAspSer 832
 DB 2661 CAAAGACCCCACTGGAGAGTGGCCCCCAGCTCTCTCT 2699

Search completed: June 22, 2003, 01:20:05
 Job time : 546 secs

THIS PAGE BLANK (USPTO)

BEST AVAILABLE COPY